А. Д. Александров
Н. Ю. Нецветаев

Геометрия

2-е издание

Допущено Научно-методическим советом по математике
Министерства образования и науки Российской Федерации
в качестве учебника для студентов высших учебных заведений,
обучающихся по направлению 010100 "Математика"

Санкт-Петербург
«БХВ-Петербург»
2010

УДК 514.1(075.8)
ББК 22.151.1я73
А46

Александров, А. Д.

А46 Геометрия: учебник / А. Д. Александров, Н. Ю. Нецветаев. — 2-е изд., исправленное. — СПб.: БХВ-Петербург, 2010. — 624 с.: ил. — (Учебная литература для вузов)

ISBN 978-5-9775-0419-5

Содержит основные разделы курса геометрии: аналитическую геометрию, элементарную геометрию на основе аксиоматики, включая геометрические преобразования и построения, элементы многомерной и проективной геометрии, дифференциальной геометрии и топологии, основания геометрии с обзором теорий «высшей» геометрии.

Для студентов высших учебных заведений

УДК 514.1(075.8)
ББК 22.151.1я73

Рецензенты:

В. Ф. Бутузов, д. ф.-м. н., профессор, заведующий кафедрой математики физического факультета Московского государственного университета (МГУ)

А. Б. Будак, к. ф.-м. н., доцент кафедры общей математики МГУ

Лицензия ИД № 02429 от 24.07.00. Подписано в печать 11.11.09.
Формат 60×90 1/16. Печать офсетная. Усл. печ. л. 39.
Тираж 2000 экз. Заказ №
"БХВ-Петербург", 190005, Санкт-Петербург, Измайловский пр., 29.

Санитарно-эпидемиологическое заключение на продукцию
№ 77.99.60.953.Д.005770.05.09 от 26.05.2009 г. выдано Федеральной службой
по надзору в сфере защиты прав потребителей и благополучия человека.

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"
199034, Санкт-Петербург, 9 линия, 12

ISBN 978-5-9775-0419-5

ОГЛАВЛЕНИЕ

Часть 2
ЭЛЕМЕНТАРНАЯ ГЕОМЕТРИЯ

Часть 3
ПРЕОБРАЗОВАНИЯ. ДРУГИЕ ГЕОМЕТРИИ

Часть 4
ДИФФЕРЕНЦИАЛЬНАЯ ГЕОМЕТРИЯ

Часть 5
ТОПОЛОГИЯ

Часть 6
ОСНОВАНИЯ ГЕОМЕТРИИ

ПРЕДИСЛОВИЕ

Настоящая книга предназначена служить пособием по курсу геометрии для студентов-математиков педагогических институтов и педагогических специальностей университетов; задача, которую она, по замыслу, должна решать, состоит в том, чтобы в наибольшей степени приблизить курс геометрии к проблемам школьного преподавания.

По содержанию книга включает материал, предусмотренный программой, и слагается из двух разделов. Первый содержит углубленное и дополненное изложение школьной геометрии в трех частях. В первой из них — аналитическая геометрия, поскольку координаты и кривые нужны в курсе анализа, а не только в самой геометрии.

Вторая часть посвящена элементарной геометрии, излагаемой на основе аксиом. В них за основной объект принята не бесконечная прямая, а отрезок — «конечная прямая», которая и фигурирует на самом деле в преподавании и в практике, как, например, в утверждении, что «прямая проводится по линейке». Понятие бесконечной прямой возникает из представления о продолжении отрезка. Равенство отрезков вводится как основное отношение, но существование длины — численной длины при данном масштабе — с условием, что равные отрезки имеют равные длины, принимается как аксиома. Аналогично дальше принимаются аксиомы о мере угла, площади и объеме (естественно, если, в частности, существование площади не доказывается, то его следует принять в качестве аксиомы).

Из аксиом выводится довольно много следствий, которые сами по себе совершенно очевидны; их вывод представляется полезным как упражнение в строгих доказательствах.

Обратим еще внимание на то необычное обстоятельство, что мы вводим в аксиоматику аксиоматическое определение понятия фигуры (или множества точек, как оно нужно в геометрии). Это представляется естественным и даже необходимым, если основывать геометрию на аксиомах, понимая под геометрией «науку о фигурах».

Какие вопросы элементарной геометрии рассматриваются вслед за аксиоматикой, читатель может видеть из оглавления.

В третьей части, тоже, собственно, относящейся к элементарной геометрии, рассматриваются отображения — в первую очередь те, которые сохраняют расстояния. Их мы называем «наложениями», как это делается в некоторых учебниках (например, у А. П. Киселева и у Л. С. Атанасяна с соавторами; удобно сказать: «одна фигура налагается на другую», а термин «движение» относится, скорее, к процессу).

За общими теоремами о наложениях и симметрией правильных многогранников следуют аффинные преобразования, инверсии и проективные преобразования, с которых начинается проективная геометрия. В нашем изложении она «сведена» к элементарной геометрии и служит тем же задачам школьного преподавания (по крайней мере в классах с углубленным изучением математики). Топологическое строение проективной плоскости выясняется потом, в части, посвященной топологии.

В конце первого раздела дается элементарное понятие о многомерной геометрии.

Второй раздел книги содержит в первых двух его частях дифференциальную геометрию и начала топологии. Они служат расширению кругозора учителя, но связаны со школьным курсом через понятия длины кривой, площади поверхности, теорему Эйлера и т. п.

Дифференциальная геометрия изучает кривые и поверхности в трехмерном пространстве (а также их обобщения в многомерном пространстве). Кривые и поверхности мы определяем как фигуры — множества точек. Мы задаем их параметрически, а не уравнениями, как в аналитической геометрии. Для краткости изложения дифференциально-геометрические понятия мы определяем чаще всего конструктивно — при помощи явной формулы, а уже затем даем чисто геометрическое определение, не использующее конкретной параметризации.

Начала теоретико-множественной топологии изложены на основе понятия топологической структуры в множестве исходя из аксиоматики открытых множеств. При этом учитывается, что первые топологические понятия для евклидова пространства вводились уже во второй части книги и, кроме того, известны читателю из курса анализа.

Порядок изложения таков, что сначала мы вводим понятие непрерывного отображения одного пространства в другое, затем — понятие гомеоморфизма и только после этого приступаем к изучению таких топологических свойств пространств и множеств в них, как связ-

ность и компактность. Изучение наиболее интересных с геометрической точки зрения пространств — многообразий — завершается топологической классификацией компактных двумерных многообразий — поверхностей.

Третью часть второго раздела представляют основания геометрии. Здесь дается аксиоматика евклидовой геометрии с теми же основными понятиями, что в первом разделе (ч. 2), но с аксиомами равенства отрезков и аксиомой непрерывности, которые дают основание для понятия длины. Соответственно, обосновываются понятия о всех величинах: длина, мера угла, площадь, объем. Дается аксиоматическое определение фигуры в элементарной геометрии, уточняющее определение, данное в ч. 2. Затем рассматриваются понятия непротиворечивости, независимости и полноты, дается модель геометрии Лобачевского и аналитическая модель евклидовой геометрии или, что равносильно, — аксиоматика, использующая координаты. Она содержит всего три аксиомы и применяется также для n-мерного пространства. Совершенно аналогичными аксиомами определяется псевдоевклидово пространство, и мы приходим к теории относительности в геометрическом ее представлении. Тут же излагается групповой принцип определения той или иной геометрии и, наконец, в дополнение, дается некоторое понятие о римановой геометрии.

Эти заключительные параграфы не претендуют на то, чтобы каждый студент мог ясно понять и усвоить то, о чем там говорится. Но, думается, в этом нет ничего страшного: учителю полезно сочетать твердое усвоение основ школьного курса с представлениями, выходящими за его пределы, как бы соединяя твердое ядро знаний с расширяющимися в пространство науки представлениями.

В книге ч. 1, 2, 3 и 6 написаны А. Д. Александровым, а 4 и 5 — Н. Ю. Нецветаевым, им же выполнено большинство рисунков.

Авторы будут благодарны всем читателям — преподавателям и студентам — за отзывы и замечания об этой книге.

Мы благодарны редактору книги А. Л. Вернеру за его большую работу по усовершенствованию рукописи и сотрудникам кафедр геометрии Воронежского и Новосибирского педагогических институтов, а также (и особенно) А. В. Кузьминых за критические замечания, которые мы по мере сил постарались учесть. Мы благодарим Т. Н. Рожковскую и Н. А. Рожковскую за помощь при оформлении рукописи.

Часть I
АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ

Основная цель первой части курса — углубить знакомство с координатами и векторами и показать, как они применяются в геометрии.

Хотя многое о координатах и векторах известно из школы, мы полностью излагаем все необходимые сведения о них. Вместе с этим рассматривается представление уравнениями в координатах линий на плоскости и поверхностей в пространстве, какие в школе не изучались.

Координаты и векторы важны не только в геометрии — они широко применяются и в других разделах математики, а также в физике и т. д. Поэтому-то речь о них и идет в самом начале нашего курса.

Существенная особенность изложения в этой книге — та, что важнейшие понятия курса часто изучаются по нескольку раз различными методами и рассматриваются с разных сторон. Так, в этой части мы трижды обращаемся к прямой на плоскости, изучая ее и координатным, и векторным методами в главах 1 и 4. Аналогично обсуждаются разные подходы к кривым второго порядка (КВП). Всестороннее изучение основных понятий призвано формировать более глубокий и широкий взгляд на них.

Глава I
НАЧАЛА АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

§ 1. Прямоугольные координаты

Аналитическая геометрия — это, коротко говоря, такой раздел геометрии, в котором геометрические фигуры изучаются с помощью алгебры на основе применения координат. Прямоугольные (декартовы) координаты на плоскости и простейшие результаты аналитической геометрии известны из школьного курса. Мы, однако, начнем с того,

что повторим их: что-что, а начало, первые основы предмета, должны быть твердо усвоены, чтобы можно было двигаться дальше.

Применение алгебры, как и последующее применение анализа, основано на представлении геометрических соотношений в виде алгебраических соотношений между числами, представляющими геометрические величины.

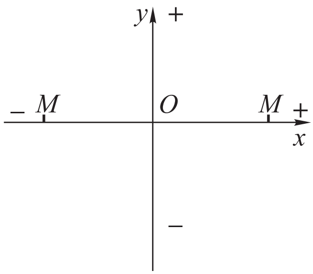


Рис. 1

Поэтому мы начинаем с того. что выбираем и закрепляем раз и навсегда[1] единицу длины (единичный отрезок) и в ней выражаем все длины и расстояния, так что дальше «длина» и «расстояние» обозначают действительные числа.

Как известно из школьного курса, прямоугольные координаты на плоскости определяются следующим образом.

Выбираем две взаимно перпендикулярные прямые — *оси координат*; их точка пересечения O — *начало координат* — делит каждую из них на два луча; один считаем положительной полуосью, другой — отрицательной (рис. 1). Каждой точке M на оси относится ее *координата*: расстояние OM, если точка на положительной полуоси, и расстояние OM с минусом для точек на отрицательной полуоси. На одной оси координата обозначается x, и эта ось называется «осью x»; на другой оси координата обозначается y, и эта ось называется «осью y» (старые названия координат: x — «абсцисса», y — «ордината» теперь мало употребимы). Для начала координат получаем $x = y = 0$.

Каждой точке M на плоскости сопоставляются в качестве ее координат x_M, y_M координаты ее проекций на оси x и y (напомним, что проекция точки на прямую — это основание опущенного из нее перпендикуляра или сама точка, если она лежит на прямой, на которую производится проектирование).

Обратно, каждой упорядоченной паре чисел x_0, y_0 соответствует точка $M(x_0,\ y_0)$ с координатами $x = x_0$, $y = y_0$. Она может быть получена следующим образом. Берем на оси x точку M_x с координатой x_0; если $y_0 = 0$, то это и будет точка $M(x_0,\ y_0)$. Если же $y_0 \neq 0$, то проводим из точки M_x перпендикуляр к оси x на длину $|y_0|$ в сторону положительных y (в полуплоскость, где лежит положительная полу-

[1]Это соглашение относится к части 1.

ось y), если $y_0 > 0$, и в сторону отрицательных y, если $y_0 < 0$. Конец этого перпендикуляра и будет точка $M(x_0,\ y_0)$. Если MM_y — опущенный из нее перпендикуляр на ось y, то координата y его основания M_y и будет y_0 (поскольку либо M на оси y, когда $x_0 = 0$, либо OM_xMM_y — прямоугольник, так что $OM_y = M_xM$ (рис. 2)). Поэтому построение точки $M(x_0,\ y_0)$ можно начинать, взяв на оси y точку M_y с координатой $y = y_0$, и из нее проводить перпендикуляр к оси y.

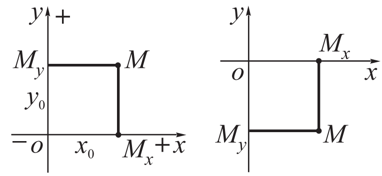


Рис. 2

Таким образом, *между точками плоскости и упорядоченными парами чисел x, y устанавливается взаимно однозначное соответствие*: каждой точке сопоставляется пара координат, и каждой паре чисел — точка с такими координатами.

Оси образуют четыре прямых угла, внутренние области которых называются *квадрантами*. В каждом квадранте координаты имеют определенный знак. Квадранты нумеруют в таком порядке: первый тот, где $x > 0$, $y > 0$; во втором $x < 0$; $y > 0$; в третьем $x < 0$, $y < 0$; в четвертом $x > 0$, $y < 0$ (рис. 3).

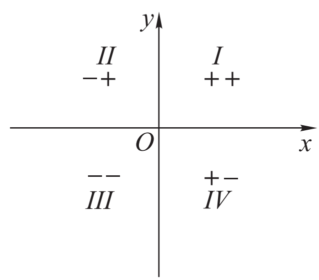


Рис. 3

Принято считать ось x горизонтальной и ограниченную ею полуплоскость, где $y > 0$, называть *верхней*, другую, где $y < 0$, — *нижней*. Направление оси x в сторону возрастания координаты x принято считать направлением вправо, противоположное — влево. Поворот от положительной полуоси x к положительной полуоси y считают идущим против часовой стрелки. Поэтому в той же последовательности — против часовой стрелки — нумерованы квадранты.

Нужно, однако, ясно понимать, что это уже не геометрия: в геометрии нет ни верха, ни низа, ни направления вправо, ни влево и никакой часовой стрелки. Все это относится не к осям координат в геометрии, а к их изображению или представлению, в частности в связи с нашим телом, по отношению к которому и определяются «вправо» и «влево».

§ 2. Прямая. Деление отрезка в данном отношении

Зафиксируем на плоскости прямоугольную систему координат x, y.

Первая задача аналитической геометрии состоит в задании или, иначе говоря, представлении фигур уравнениями.

Определение. Уравнение с двумя переменными x, y *задает* фигуру F или, как говорят, является *уравнением* этой фигуры в *прямоугольных координатах*, если

а) координаты всякой точки фигуры F удовлетворяют уравнению;

б) всякая пара чисел x, y, удовлетворяющих уравнению, оказывается парой координат некоторой точки фигуры F.

Другими словами, точка принадлежит фигуре тогда и только тогда, когда координаты точки удовлетворяют уравнению фигуры. При этом второе условие допускает полезную переформулировку:

б$'$) если точка не принадлежит фигуре, то ее координаты не удовлетворяют уравнению последней.

Определению уравнения фигуры соответствует определение того, что значит «фигура, заданная или определенная данным уравнением»:

Фигура, определенная данным уравнением в некоторой данной системе координат, — это множество всех точек плоскости, координаты которых удовлетворяют этому уравнению.

Говоря об уравнении с двумя переменными, не исключают, что одна из переменных может в нем отсутствовать. Тогда уравнение не налагает на нее никаких условий, и она может независимо принимать любые значения.

Определение. Уравнение вида

$$ax + by + c = 0 \tag{1}$$

называется *уравнением* ***первой степени*** или ***линейным,*** если хотя бы один из коэффициентов a, b отличен от нуля $(a^2 + b^2 \neq 0)$.

Теорема 1. *Всякая прямая задается в прямоугольных координатах линейным уравнением, и обратно: всякое линейное уравнение задает в прямоугольных координатах прямую.*

Доказательство. Докажем первую часть теоремы. Пусть на плоскости введены прямоугольные координаты и дана какая-либо прямая. Если она параллельна оси y, то ее уравнение имеет вид $x = a$ $(a = \text{const})$[2].

[2] Так обозначают постоянную величину (от лат. constans — постоянный).

Действительно, такая прямая перпендикулярна оси x. Поэтому у ее точек $x = x_0$, где x_0 — координата точки ее пересечения с осью x (как это следует из самого определения координат). Для всех же других точек $x \neq x_0$ (рис. 4).

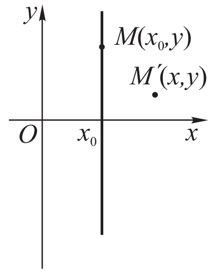


Рис. 4

Совершенно аналогично прямая, параллельная оси x, задается уравнением вида $y = a$ ($a = \text{const}$).

Пусть теперь данная прямая проходит через начало координат и не совпадает ни с той, ни с другой осью. Тогда она пересекает в начале координат ось x, и ее луч, расположенный в верхней полуплоскости, образует с положительной полуосью x некоторый угол α, отличный от прямого (рис. 5).

В таком случае для каждой точки прямой имеем

$$\frac{y}{x} = \operatorname{tg}\alpha,$$

как это следует из самого определения тангенса (для луча, расположенного в нижней полуплоскости, x и y — противоположных знаков, а потому отношение их то же). Для точек, не лежащих на данной прямой, отношение другое.

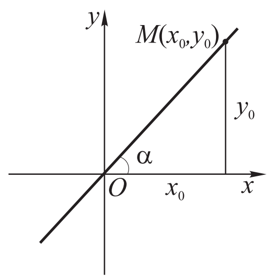


Рис. 5

Таким образом, прямая задается уравнением вида

$$y = kx \quad (k = \operatorname{tg}\alpha). \qquad (2)$$

Пусть теперь прямая p не проходит через начало, а пересекает ось y в точке A с координатой $y = b$, так что $b = \pm OA$. Проведем через начало прямую q, параллельную p (рис. 6). Каждая прямая, параллельная оси y, пересекает обе прямые в некоторых точках P, Q, и мы получаем параллелограмм $OAPQ$, так что $QP = OA$. Поэтому координата y точки P отличается от координаты точки Q на столько же, на сколько координата A отличается

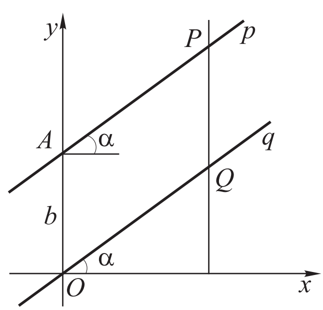


Рис. 6

от координаты O, т. е.

$$y_P - y_Q = y_A - y_O, \quad y_P = y_Q + b.$$

Прямая q проходит через начало. Поэтому на ней

$$y = kx \quad (y_Q = kx_Q).$$

Поэтому на прямой p выполняется равенство

$$y = kx + b. \tag{3}$$

Ни для какой точки, не лежащей на прямой p, это равенство не выполняется, так как «отступление по вертикали» от прямой q, т. е. $y - kx$, будет другим.

Число k называется ***угловым коэффициентом*** прямой p.

Итак, уравнение (3) представляет прямую p.

Таким образом, мы получили следующий результат.

Всякая прямая, не параллельная оси y, представляется уравнением вида

$$y = kx + b,$$

где ***угловой коэффициент*** $k = \operatorname{tg} \alpha$ *— тангенс угла наклона прямой к оси x и b — координата точки пересечения прямой с осью y.*

Прямая, параллельная оси y, представляется уравнением $x = a$.

Докажем теперь, что всякое линейное уравнение представляет прямую.

Пусть дано уравнение вида (1) с $a^2 + b^2 \neq 0$. Допустим, что в нем $b = 0$. Тогда $a \neq 0$, и уравнение приводится к виду $x = -c/a$, т. е. оно представляет прямую, параллельную оси y. Если же $b \neq 0$, то уравнение приводится к виду

$$y = -\frac{a}{b}x - \frac{c}{b}, \quad \text{т. е.} \quad y = kx + d.$$

Оно, следовательно, представляет прямую (с угловым коэффициентом k, проходящую через точку $(0,\ d)$).

Итак, теорема 1 доказана полностью. $\square$

Дополнение к теореме 1. *Два линейных уравнения вида* (1) *представляют одну и ту же прямую тогда и только тогда, когда их левые части отличаются только множителем, т. е. когда они имеют вид*

$$ax + by + c = 0, \quad \lambda ax + \lambda by + \lambda c = 0, \quad \lambda \neq 0.$$

Доказательство. То, что уравнения такого вида представляют одну и ту же прямую, очевидно: достаточно разделить второе уравнение на λ.

Покажем обратное. Пусть уравнения

$$ax + by + c = 0, \quad a_1x + b_1y + c_1 = 0$$

представляют одну и ту же прямую. Если в одном $b = 0$, то оно представляет прямую, параллельную оси y, значит другое — тоже, и в нем тоже должно быть $b_1 = 0$. Поэтому они приводятся к виду

$$x = -\frac{c}{a}, \quad x = -\frac{c_1}{a_1},$$

так что $c/a = c_1/a_1$, т. е. $a_1 : a = c_1 : c$. Осталось положить $\lambda = a_1/a$.

Если же $b \neq 0$ и $b_1 \neq 0$, то уравнения приводятся к виду

$$y = -\frac{a}{b}x - \frac{c}{b}, \quad y = -\frac{a_1}{b_1}x - \frac{c_1}{b_1}.$$

Но они представляют одну прямую, только если совпадают угловые коэффициенты и свободные члены, т. е. если $a_1/b_1 = a/b$, $c_1/b_1 = c/b$. Полагая $b_1/b = \lambda$, имеем $a_1 = \lambda a$, $b_1 = \lambda b$, $c_1 = \lambda c$, что и требовалось доказать. $\square$

Деление отрезка в данном отношении.

Теорема 2. *Середина отрезка с концами* $A_1(x_1,\ y_1)$, $A_2(x_2,\ y_2)$ *есть точка с координатами*

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}.$$

Доказательство. Проекция середины отрезка есть, очевидно, середина его проекции. Середина же отрезка оси между точками с координатами x_1, x_2 и есть точка с координатой $(x_1 + x_2)/2$ (так как в середине должно быть $x - x_1 = x_2 - x$; рис. 7). То же верно и для y.

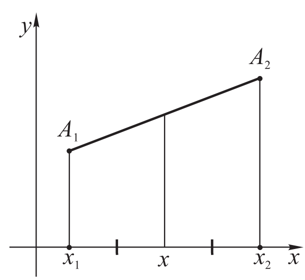


Рис. 7

Обобщение. Если точка A_t на отрезке A_0A_1 такова, что $A_0A_t = A_0A_1 \cdot t$, то

$$x_t = (1 - t)x_0 + tx_1,$$
$$y_t = (1 - t)y_0 + ty_1.$$

(Докажите это.)

§ 3. Расстояние между точками. Окружность. Прямая

В формулировках теорем 1–3 фиксируем прямоугольные координаты x, y.

Теорема 1. *Расстояние между точками* $A(x_1,\ y_1)$, $B(x_2,\ y_2)$ *выражается формулой*

$$AB = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \tag{1}$$

Доказательство. Если точки A, B лежат на оси x, то расстояние между ними равно $AB = |x_2 - x_1|$. Если $y_1 = y_2 \neq 0$, то отрезок AB параллелен оси x, и потому $AB = A'B'$, где A', B' — проекции точек A, B на ось x. Поэтому $AB = |x_2 - x_1|$. Это совпадает с (1) при $y_1 = y_2$ (рис. 8).

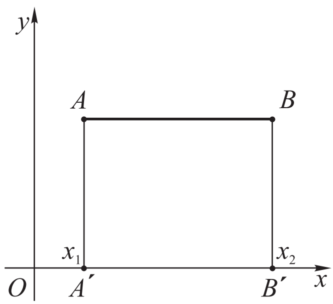


Рис. 8

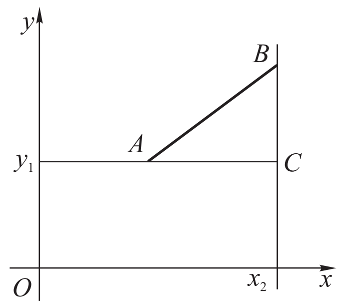


Рис. 9

Совершенно так же при $x_1 = x_2$ получаем $AB = |y_2 - y_1|$ — то же, что дает (1).

Пусть теперь $x_1 \neq x_2$ и $y_1 \neq y_2$. Проведем через A прямую, параллельную оси x, а через B — прямую, параллельную оси y; это прямые, на которых $y = y_1$ и $x = x_2$. Они взаимно перпендикулярны, так как параллельны осям, и пересекаются в точке $C(x_2,\ y_1)$ (рис. 9). По предыдущему выводу мы имеем

$$AC = |x_1 - x_2|, \quad BC = |y_1 - y_2|.$$

Отрезки AC, BC, AB образуют прямоугольный треугольник, и по теореме Пифагора

$$AB^2 = AC^2 + BC^2 = |x_1 - x_2|^2 + |y_1 - y_2|^2,$$

что равносильно равенству (1). $\square$

Теорема 2. *Окружность с центром $A(x_0,\ y_0)$ и радиусом r представляется уравнением*

$$(x - x_0)^2 + (y - y_0)^2 = r^2. \tag{2}$$

Доказательство. Это непосредственно следует из определения окружности и теоремы 1. Действительно, точка с координатами x, y лежит на окружности тогда и только тогда, когда расстояние от нее до точки A равно r, т. е. когда x и y удовлетворяют уравнению (2). □

Уравнение (2) приводится к виду

$$x^2 + y^2 + ax + by + c = 0, \tag{3}$$

где $a = -2x_0$, $b = -2y_0$, $c = x_0^2 + y_0^2 - r^2$.

Естественно спросить, что вообще может представлять уравнение вида (3).

Теорема 3. *Уравнение* (3) *представляет одну из трех фигур:*

а) *окружность, если* $a^2 + b^2 - 4c > 0$,

б) *одну точку, если* $a^2 + b^2 - 4c = 0$,

в) *пустое множество*[3], если $a^2 + b^2 - 4c < 0$.

Доказательство. Положим в уравнении (3)

$$a = -2x_0, \quad b = -2y_0. \tag{4}$$

Тогда его можно переписать в виде

$$(x - x_0)^2 + (y - y_0)^2 - x_0^2 - y_0^2 + c = 0$$

или, пользуясь (4) и перенося свободные члены вправо, в виде

$$(x - x_0)^2 + (y - y_0)^2 = \frac{1}{4}(a^2 + b^2 - 4c). \tag{5}$$

Если

$$a^2 + b^2 - 4c > 0, \tag{6}$$

то можно обозначить в (5) правую часть через r^2, и мы получим уравнение (2). Итак, при выполнении неравенства (6) уравнение (3) представляет окружность с центром в точке $(x_0,\ y_0)$ и радиусом r, где

$$x_0 = -\frac{a}{2}, \quad y_0 = -\frac{b}{2}, \quad r = \frac{1}{2}\sqrt{a^2 + b^2 - 4c}.$$

[3]Пустое множество — «ничего», т. е. в случае (3) уравнение «ничего не представляет».

Если правая часть в (5) равна нулю, то получаем

$$(x-x_0)^2+(y-y_0)^2=0.$$

Это уравнение удовлетворяется только при $x=x_0$, $y=y_0$, т. е. оно представляет точку $(x_0,\ y_0)$.

Наконец, если правая часть в (5) отрицательна, то равенство (5) невозможно, так как левая его часть заведомо неотрицательна при всех x, y. $\square$

Рассмотрев уравнение (3), мы получили пример ко ***второй задаче*** аналитической геометрии: исследованию того, какие фигуры может представлять то или иное уравнение.

Второй вывод уравнения прямой. Как известно, множество точек, каждая из которых равноудалена от двух данных точек, представляет собой прямую. Очевидно, всякую прямую можно представить как такое множество — как серединный перпендикуляр подходящего отрезка. Это позволяет получить уравнение прямой другим способом.

Пусть выбраны прямоугольные координаты x, y и дана какая-либо прямая a. Проведем перпендикулярный ей отрезок A_1A_2, который она делит пополам. Тогда точки M прямой a характеризуются равенством расстояний: $A_1M=A_2M$ или $A_1M^2=A_2M^2$. Если координаты точек A_1, A_2 обозначить a_1, b_1; a_2, b_2, а координаты произвольной точки $M - x,\ y$, то, применяя формулу (1) для расстояния, можно написать:

$$(x-a_1)^2+(y-b_1)^2=(x-a_2)^2+(y-b_2)^2. \tag{7}$$

Выполняя возведение в квадрат, заметим, что с обеих сторон равенства оказывается x^2+y^2. Поэтому квадраты взаимно уничтожаются и остаются x, y только в первой степени, т. е. получаем линейное уравнение

$$ax+by+c=0 \tag{8}$$

(как выражаются здесь a, b, c через a_1, a_2; b_1, b_2, нам пока не важно).

Мы еще раз доказали, таким образом, что всякая прямая представляется линейным уравнением.

Но на сделанный вывод можно посмотреть и иначе. Мы доказали, что множество точек M, для которых $A_1M=A_2M$, представляется линейным уравнением. А так как раньше было доказано, что линейное уравнение задает прямую, то мы, можно сказать, доказали теорему геометрии: множество точек, равноудаленных от двух данных точек, представляет собой прямую.

Теперь докажем, что всякое линейное уравнение задает прямую. Для этого выпишем выражения коэффициентов a, b, c в (8), как они получаются из (7):

$$\begin{gathered} a = 2(a_2 - a_1), \quad b = 2(b_2 - b_1), \\ c = (a_1^2 + b_1^2) - (a_2^2 + b_2^2). \end{gathered} \tag{9}$$

Если дано уравнение $ax+by+c=0$, то подберем a_1, b_1; a_2, b_2 так, чтобы удовлетворить равенства (9), и тогда убедимся, что данное уравнение представляет множество точек, равноудаленных от точек $A_1(a_1,\ b_1)$, $A_2(a_2,\ b_2)$ — т. е. представляет прямую. (Проведите этот вывод, решив уравнения (9).)

Окружность Аполлония. В задачи аналитической геометрии входит получение геометрических результатов при помощи координат и, соответственно, алгебры.

Прекрасный тому пример дает задача: пусть A, B — две данные точки. Спрашивается, какую фигуру образуют такие точки M, для которых отношение расстояний AM, BM одно и то же

$$\frac{AM}{BM} = k = \text{const?} \tag{10}$$

При $k=1$ это будет, как известно, прямая. Но при $k \neq 1$ это — окружность; ее называют *окружностью Аполлония*[4].

Для доказательства возьмем прямоугольные координаты с началом в точке B и положительной полуосью x, проходящей через A. Тогда, полагая $AB = a$, имеем (рис. 10)

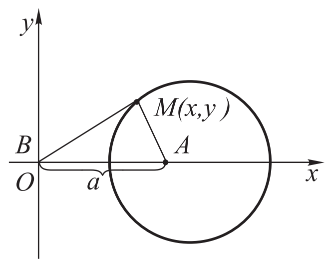


Рис. 10

$$\begin{gathered} BM^2 = x^2 + y^2, \\ AM^2 = (x-a)^2 + y^2. \end{gathered}$$

А по условию:

$$AM^2 = k^2 \cdot BM^2.$$

Очевидно, мы получим уравнение вида (3), а из него найдем центр и радиус окружности. (Проведите это решение в деталях.)

[4]Аполлоний (около 262–190 до н. э.) — один из великих древнегреческих геометров эпохи Евклида (конец 4 в. — 3 в. до н. э.) и Архимеда (ок. 287–212 до н. э.).

Получить решение этой задачи, не пользуясь координатами, затруднительно.

Обобщение. Пусть дано произвольное число n точек $A_1, \ldots, A_n$ и числа $k_1, \ldots, k_n$. Спрашивается, что представляет собою множество точек M, для которых

$$k_1 \cdot A_1M^2 + k_2 \cdot A_2M^2 + \ldots + k_n \cdot A_nM^2 = c = \text{const}? \qquad (11)$$

(Условие (10) задачи Аполлония — частный случай: в нем точек две, $c = 0$, $K_1 = 1$, $k_2 = -k^2$.) Ответ: это будет либо окружность, либо прямая, либо одна точка, либо пустое множество (как в случае уравнения (3)). То же получится, если в (11) справа поставить $ax + by + c$. (Докажите все это.)

§ 4. Полярные и другие координаты

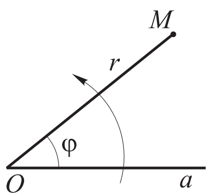


Рис. 11

Выберем точку O, исходящий из нее луч a и направление отсчета углов от луча a вокруг точки O. Каждой точке плоскости M, отличной от O, сопоставляем два числа: расстояние r от O до M и выраженный в радианах угол φ, образуемый лучом OM с лучом a и отсчитываемый в выбранном направлении (рис. 11) (в противоположном направлении угол считается отрицательным). Эти числа r, φ называются *полярными координатами* точки M; точка O — *центр*, луч a — *начальный луч* системы полярных координат. Центру O соответствует $r = 0$, а угол φ не определен.

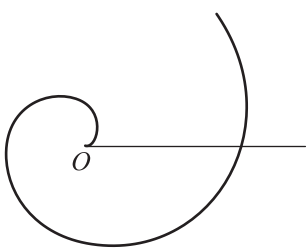


Рис. 12

Для угловой координаты — для угла φ — допускаются, вообще говоря, любые значения; при этом, разумеется, углы, отличающиеся на целое кратное 2π, определяют один и тот же луч с началом O.

Таким образом, в отличие от прямоугольных координат, между значениями полярных координат и точками плоскости нет взаимно однозначного соответствия: каждая пара координат задает определенную точку, но одной точке соответствует одно расстояние r и бесконечно много значений

угловой координаты. Изменение координаты φ подвижной точки M удобно представлять как вращение луча OM. Это видно на примере *спирали Архимеда* (рис. 12), задаваемой уравнением

$$r = a\varphi \quad (a > 0).$$

Каждый луч из O пересекает эту спираль в бесконечном числе точек со значениями φ, отличающимися на целое кратное 2π (так как $r \geqslant 0$, то $\varphi \geqslant 0$).

О координатах. *Координатами* (точнее координатами точек) вообще называются числа, которые задают точки, — точка фиксируется указанием ее координат. Координаты или, как говорят, системы координат могут быть самые разные: прямоугольные и полярные — только самые простые; кроме этих мыслимы и применяются другие, некоторые из них указаны дальше. Кроме того, сами прямоугольные координаты можно вводить, беря разные оси, так же как полярные — беря разные начала, начальные лучи и меняя направление отсчета углов.

В связи с разнообразием возможных систем координат нужно иметь в виду следующее. Когда говорят о задании фигуры уравнением, нужно оговорить — в каких координатах. Например: что задает уравнение $u = av$ в координатах u, v? Если координаты u, v прямоугольные, то — прямую; если полярные — то спираль Архимеда. Обычно опускают такие оговорки о системе координат, когда ясно, какие координаты имеются в виду, чаще всего — прямоугольные.

Аффинные координаты. Оси x, y (каждая — со своим масштабом!) проводятся под произвольным углом; угол между положительными полуосями пусть будет α. Координата x точки M — это координата той точки на оси x, в которой эту ось пересекает проведенная через M прямая, параллельная оси y. Аналогично определяется координата y (рис. 13).

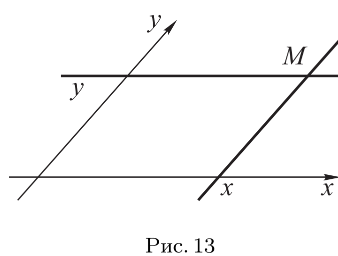


Рис. 13

Докажите — в таких координатах — теорему о прямой: всякая прямая представляется линейным уравнением, и обратно: всякое линейное уравнение представляет прямую.

Найдите выражение для расстояния между точками. Напишите уравнение окружности.

§ 5. Преобразование координат

Переход от одних координат к другим — от одной системы координат к другой — совершается путем *преобразования координат.* Формулы преобразования выражают старые координаты через новые. (Координаты, от которых происходит переход, можно называть «старыми», а другие — «новыми».) Рассмотрим переход от прямоугольных координат к полярным и обратно, когда у них начало общее, положительная полуось x совпадает с начальным лучом и направление отсчета углов идет от положительной полуоси x к положительной полуоси y. Тогда (как можно видеть из самого определения тех и других координат, а также — косинуса и синуса (рис. 14))

$$\begin{gathered} x = r\cos\varphi, \quad y = r\sin\varphi; \\ r = \sqrt{x^2+y^2}, \quad \operatorname{tg}\varphi = \frac{y}{x}. \end{gathered} \tag{1}$$

(Так как в пределах от 0 до 2π угол не определяется тангенсом однозначно, то нужно учитывать знак x и y. Как?)

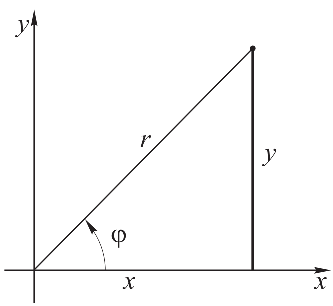


Рис. 14

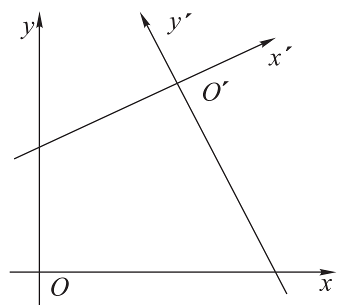


Рис. 15

Рассмотрим теперь преобразование прямоугольных координат. Пусть имеются две системы координат: старая x, y и новая x', y'. Начало старой — O, новой — O' (рис. 15). Геометрически переход от старой системы координат к новой можно представить таким образом: переносим оси так, чтобы начало O попало в O'. Затем поворачиваем оси так, чтобы положительная полуось x перешла в положительную полуось x'. После этого положительная полуось y либо совпадает с положительной полуосью y', либо не совпадает, а будет направлена противоположно. Тогда надо ее «перевернуть» отражением относительно оси x.

Представим это формулами в координатах. Рассмотрим каждый вид преобразования координат отдельно.

Перенос начала. Пусть новое начало O' имеет старые координаты x_0, y_0 (рис. 16). Новые оси параллельны старым и одинаково с ними направлены. Эти оси — прямые $x = x_0$, $y = y_0$. Поэтому новые координаты x', y' произвольной точки $M(x,\ y)$ будут

$$x' = x - x_0, \quad y' = y - y_0. \tag{2}$$

Поворот вокруг начала. Пусть новые координаты имеют то же начало O, что и старые, но оси повернуты на угол α (отсчитываемый как было указано) (рис. 17).

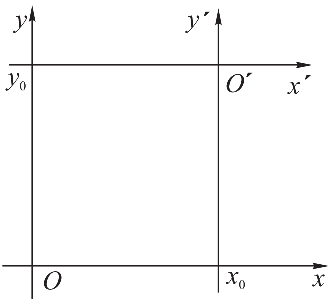


Рис. 16

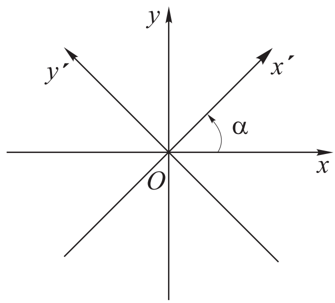


Рис. 17

Если ввести полярные координаты со старым и новым начальным лучом, то новые полярные координаты будут, очевидно,

$$r' = r, \quad \varphi' = \varphi - \alpha. \tag{3}$$

Поэтому из формул (1) новые x', y' будут

$$x' = r' \cos\varphi' = r\,\cos(\varphi - \alpha),$$
$$y' = r'\,\sin\varphi' = r\,\sin(\varphi - \alpha).$$

Отсюда

$$x' = r\,\cos\varphi\,\cos\alpha + r\,\sin\varphi\,\sin\alpha,$$
$$y' = r\,\sin\varphi\,\cos\alpha - \cos\varphi\,\sin\alpha.$$

И ввиду выражений (1) для старых x, y получаем

$$x' = x\,\cos\alpha + y\,\sin\alpha, \quad y' = y\,\cos\alpha - x\,\sin\alpha. \tag{4}$$

Обратный переход от новых координат x', y' к старым x, y сводится к повороту на угол α в обратную сторону, т. е. на $-\alpha$. Поэтому (так как $\cos(-\alpha) = \cos\alpha$, $\sin(-\alpha) = -\sin\alpha$)

$$x = x' \cos\alpha - y' \sin\alpha, \quad y = x' \sin\alpha + y' \cos\alpha. \tag{5}$$

Если соединить поворот с переносом, который делается сначала, то в формулу (4) надо подставить $x - x_0$, $y - y_0$ вместо x, y. Это приводит к тому, что формулы (4) получат свободный член. Аналогично будет с формулами (5). Выписывать эти формулы не будем.

Изменение направления оси y означает перемену знака координаты y, так что формулы преобразования будут

$$x' = x, \quad y' = -y.$$

Между прочим, отсюда следует полезное замечание: если уравнение $f(x, y) = 0$ не изменяется при перемене знака y, то оно определяет фигуру, симметричную относительно оси x.

§ 6. Об аналитической геометрии

Как уже было сказано в § 1, общая задача аналитической геометрии состоит в изучении геометрических фигур с помощью алгебры на основе применения координат. Для этого фигура представляется уравнениями или неравенствами в координатах — в таком представлении состоит так называемая *первая задача* аналитической геометрии.

В аналитическую геометрию включается также *вторая задача,* обратная первой: исследовать, какие геометрические фигуры представятся теми или иными уравнениями. Все это можно было видеть в предшествующем изложении, особенно в § 3, где мы вывели уравнение окружности в прямоугольных координатах, затем выяснили, какие фигуры может вообще представлять уравнение подобного вида, затем то же сделали для прямой и, наконец, дали решение геометрической задачи об окружности Аполлония, пользуясь результатами, полученными в координатах.

Хотя мы привели примеры разных систем координат, преимущественно используют прямоугольные координаты.

Уравнения. *Уравнением с двумя переменными* x, y в общем смысле называют равенство $F(x, y) = 0$, где $F(x, y)$ — вообще какая-либо функция переменных вещественных чисел x, y. Уравнение называется ***алгебраическим*** *уравнением степени* n, если

его левая часть представляет собой многочлен степени n относительно x, y с численными коэффициентами. Это будут уравнения вида $ax + by + c = 0$ (первой степени), $ax^2 + bxy + cy^2 + dx + ey + f = 0$ (второй степени) и т. д. При этом подразумевается, что хотя бы при одном из «старших» членов коэффициент не равен нулю. Но, например, члены с y могут вовсе отсутствовать. Поэтому число переменных определяется не тем, сколько их в выражении левой части уравнения, а условием, что уравнение относится к *двум* переменным: левая часть $F(x,\ y)$ есть функция *двух* переменных. Если одна переменная в уравнение не входит, то ее значения ничем не связаны. (По этому поводу заметим, что определять левую часть уравнения $F(x,\ y) = 0$ как «выражение, содержащее x и y», неправильно; к тому же, если y отсутствует, его можно «включить», прибавив $y - y$.)

В пространстве — три координаты, и, соответственно, там уравнение $F(x,\ y) = 0$ имеет совсем другой смысл, чем в случае двух переменных — на плоскости.

В аналитической геометрии принято говорить, что алгебраическое уравнение $F(x,\ y) = 0$ задает *линию* или *кривую*, что одно и то же (прямая — тоже кривая). Но мы уже видели, что уравнение может задавать и точку, и пустое множество. В аналитической геометрии это тоже «кривые» или «линии». В других частях математики такое понимание термина «кривая» исключается, а термин «линия» не употребляется.

В случае алгебраических кривых двум определениям — уравнения данной кривой и кривой, заданной уравнением, — соответствуют все те же две задачи: вывести уравнение данного вида кривых и, обратно, выяснить, какие «кривые» могут определяться уравнениями данного вида. Эти задачи мы решили в § 2 для уравнений первой степени (в прямоугольных координатах): они задают прямые, и каждая прямая может быть задана таким уравнением. В § 3 те же две задачи были решены для уравнений второй степени вида $x^2 + y^2 + ax + by + c = 0$. В следующей главе вторая задача будет решена для любых уравнений степени 2.

Если две кривые задаются уравнениями

$$F(x, y) = 0, \quad G(x,\ y) = 0,$$

то их объединение задается уравнением

$$F(x,\ y)\ G(x,\ y) = 0,$$

а пересечение — уравнением

$$F^2(x,\ y) + G^2(x,\ y) = 0$$

или системой уравнений

$$\begin{cases} F(x,\ y) = 0, \\ G(x,\ y) = 0. \end{cases}$$

Так, между прочим, можно из двух прямых получить кривые второго порядка (чтобы выяснить какие — рассмотрите возможные случаи). Кроме того, можно заметить, что кривая, задаваемая уравнением $F(x,\ y) = 0$, задается также уравнением $F^2(x,\ y) = 0$. Поэтому, кстати, прямая — это кривая не только первого, но и второго порядка.

Это замечание, как и другие, обращает внимание на то, что нужно быть по возможности точным, но не догматичным (а догматизм в школьном преподавании следует преодолевать).

Неравенства. Фигуры, особенно имеющие внутренние точки, задают в координатах одним или несколькими неравенствами, — как говорят, *системой* неравенств. Например, в прямоугольных координатах неравенство

$$x^2 + y^2 \leqslant r^2$$

задает круг радиуса r с центром в начале. А система из двух неравенств

$$\{x^2 + y^2 \leqslant r^2,\ y \geqslant 0\}$$

задает полукруг. Неравенство $x^2 + y^2 < r^2$ задает внутренность круга. Неравенство

$$ax + by + c \leqslant 0 \quad (ax + by + c \geqslant 0)$$

задает полуплоскость, ограниченную прямой $ax+by+c = 0$ (убедитесь в этом).

Общее определение неравенства (или системы неравенств), задающего данную фигуру, повторяет определение уравнения фигуры, и точно так же определение фигуры, заданной неравенством, повторяет определение фигуры, заданной уравнением. (Сформулируйте соответствующие определения!)

Параметрическое задание кривых. *Параметрическим заданием* кривой называется такое ее задание, когда она представляется как путь (траектория) движения точки. То есть, выражаясь математически, можно сказать так: параметрическое задание кривой состоит в

том, что координаты ее точек задаются как функции некоторой переменной — «параметра»:

$$\{x = f(t),\ y = g(t)\}.$$

Каждому значению параметра $t = t_0$ из указанного промежутка его изменения соответствует точка с координатами $x = f(t_0)$, $y = g(t_0)$, и так представляются все точки кривой. Если представлять параметр как время, то и получается сказанное выше о кривой, как о пути (рис. 18).

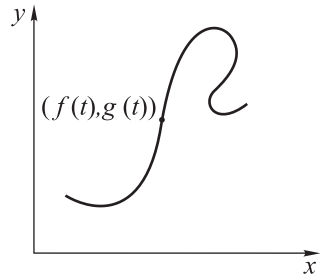


Рис. 18

Пример: параметрическое представление окружности

$$x = r\ \cos t, \quad y = r\ \sin t,$$

где t «пробегает» интервал от 0 до 2π (или, если угодно, от $-\infty$ до $+\infty$). Очевидно, два эти уравнения равносильны тому, что $x^2 + y^2 = r^2$.

Параметрическое представление прямой:

$$x = at + b, \quad y = ct + d \quad (-\infty < t < +\infty),$$

где $a^2 + c^2 \neq 0$ (хотя бы одно из чисел a, c отлично от нуля). Такие уравнения задают прямую, и всякую прямую можно представить такими уравнениями. (Убедитесь в этом!) Впрочем, дальше мы специально рассмотрим параметрическое задание прямой (см. § IV.4).

В аналитической геометрии параметрическим заданием кривых пользуются мало, но в общей теории кривых параметрическое задание является основным. Эта общая теория кривых принадлежит *дифференциальной геометрии,* называемой так потому, что в ней важнейшую роль играет применение дифференциального исчисления (см. ч. 4, гл. I).

Геометрия и анализ. В анализе изучают функции $f(x)$ или, что равносильно, их графики, заданные уравнениями $y = f(x)$. Тут функция изображается кривой. Но не сама по себе, а в отношении к данной системе координат. В других координатах x', y' та же кривая будет задаваться другой функцией или, вообще, некоторым уравнением $F(x',\ y') = 0$.

Можно сказать так. Анализ изучает кривые в связи с данной системой координат, в которой они представляются уравнением вида

$y = f(x)$. Геометрия изучает кривые независимо от той или иной системы координат. Геометрические свойства кривой не зависят от выбора системы координат, в которой кривая представляется уравнением. К геометрии относится то, что не изменяется при преобразовании координат. То, что не изменяется, называется (с латинского) *инвариантным.* Поэтому говорят: предмет геометрии составляют *инварианты* преобразований прямоугольных координат.

Глава II

КРИВЫЕ ВТОРОГО ПОРЯДКА

Здесь, говоря о координатах, мы будем иметь в виду прямоугольные координаты (пока явно не обратимся к полярным координатам).

§ 1. Типы кривых второго порядка

Мы знаем, что уравнения первой степени представляют прямые линии. Естественно задаться вопросом о множествах, определяемых в прямоугольных координатах уравнениями второй степени. Такое общее уравнение имеет вид

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_1x + 2a_2y + a_0 = 0; \tag{1}$$

предполагается, конечно, что хотя бы один из коэффициентов a_{11}, a_{12}, a_{22} отличен от нуля[5].

Определение. Геометрическое место (или множество) точек, координаты которых удовлетворяют такому уравнению, называется *кривой второго порядка* (сокращенно ***КВП***).

Это определение инвариантно относительно выбора системы координат: координаты в одной системе выражаются через координаты в другой системе линейно, так что подстановка таких выражений для x и y в уравнение (1) не может повышать его степень (см. § 7). При этом члены второй степени не могут и исчезнуть. Иначе при обратном переходе от «новых» координат x', y' к «старым» x, y они не могли бы появиться.

Оказывается, уравнения второй степени могут представлять существенно разные кривые. *Существует восемь типов кривых второго порядка.* Перечислим их.

---

[5]Принято писать $2a_{12}$, $2a_1$, $2a_2$ для удобства, которое скажется дальше; $2a_{12}xy = a_{12}xy + a_{21}yx$, $2a_1x = a_1x \cdot 1 + a_1\, 1 \cdot x$, ...

Во-первых, это *эллипсы, гиперболы и параболы;* формально их можно определить так.

1. ***Эллипс*** — это кривая (рис. 19), представляемая в подходящей системе координат уравнением вида

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{2}$$

(Как правило, подразумевают, что $a \geqslant b > 0$.) Простейший случай — когда $a = b$. Тогда уравнение представляет окружность

$$x^2 + y^2 = a^2.$$

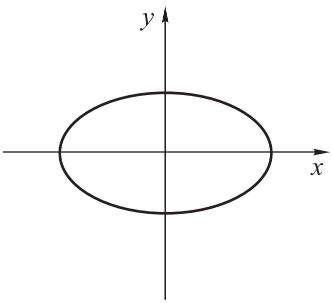


Рис. 19

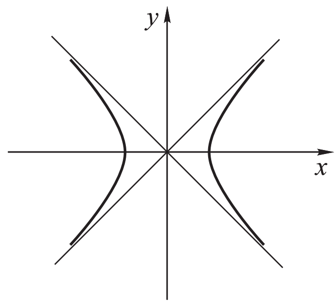


Рис. 20

Таким образом, окружность — простейший частный случай эллипса.

2. ***Гипербола*** — это кривая (рис. 20), которая в подходящих координатах представляется уравнением вида

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{3}$$

Простейший частный случай получается при $a = b$, — это так называемая *равнобокая* гипербола (рис. 21), представляемая уравнением

$$x^2 - y^2 = a^2.$$

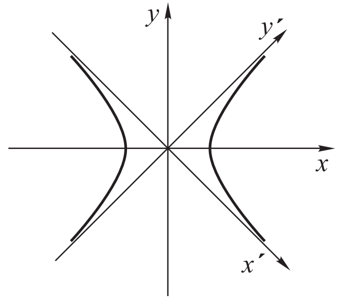


Рис. 21

Если повернуть оси на 45°, это уравнение преобразуется в

$$2x'y' = a^2.$$

Так что в таких координатах равнобокая гипербола представляет график обратной пропорциональности.

3. ***Парабола*** — это кривая (рис. 22), которая в подходящих координатах представляет график пропорциональности квадрату или, другими словами, график простейшей квадратной функции

$$y = ax^2.$$

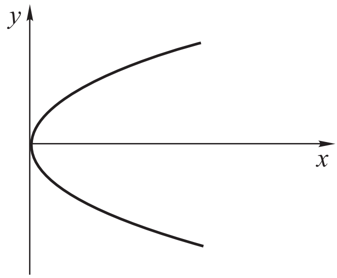


Рис. 22

Однако в аналитической геометрии принято менять оси ролями и задавать параболу уравнением

$$y^2 = 2px. \tag{4}$$

Почему именно так — будет видно дальше (см. § 5). Уравнения (2), (3), (4), какими мы определили здесь эллипсы, гиперболы и параболы, называются их ***каноническими уравнениями*** (понятно, что в других координатах они будут задаваться уравнениями другого вида).

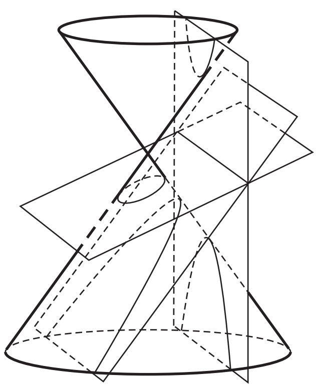


Рис. 23

Кривые трех названных типов — эллипсы, гиперболы, параболы — обладают рядом замечательных геометрических свойств и могут быть определены разными геометрическими условиями; это будет сделано дальше (см. § 3, 4, 5). В частности, они представляют собою кривые, получающиеся, если прямой круговой конус (конус вращения), образуемый прямыми, проходящими через одну точку — вершину конуса, пересекать плоскостями, не проходящими через вершину (рис. 23). Поэтому они называются ***коническими сечениями.***

Их можно было бы назвать также «настоящими» КВП в сравнении с остальными пятью типами КВП.

За коническими сечениями следуют три типа КВП, состоящих из прямых.

4. ***Пара пересекающихся прямых*** (рис. 24). В подходящих координатах она представляется уравнением $y^2 - k^2x^2 = 0$.

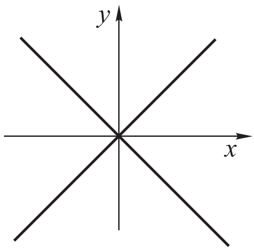


Рис. 24

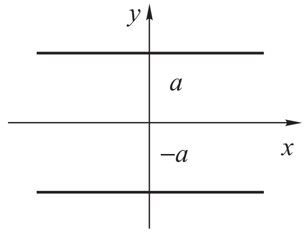


Рис. 25

5. ***Пара параллельных прямых*** (рис. 25). В подходящих координатах она представляется уравнением $y^2 - a^2 = 0$.

6. ***Одна прямая*** (рис. 26). В подходящих координатах она представляется уравнением $y^2 = 0$.

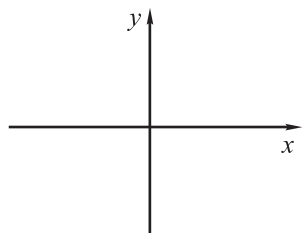


Рис. 26

Наконец, имеются еще два типа вырождающихся КВП:

7. ***Одна точка.*** Ее представляет уравнение $x^2 + y^2 = 0$ в таких координатах, в которых она принята за начало.

8. ***Пустое множество*** — пустая «кривая» без точек. Ее представляет уравнение типа

$$x^2 + y^2 + 1 = 0 \quad \text{или} \quad x^2 + 1 = 0.$$

С этими КВП мы уже встречались, рассматривая обобщение уравнения окружности в § 3 гл. I.

Итого мы перечислили 8 типов КВП. Других КВП не бывает, как будет доказано дальше в § 7. А теперь рассмотрим некоторые геометрические свойства конических сечений.

§ 2. Форма эллипса, гиперболы и параболы

Исследуем форму эллипса, гиперболы и параболы, исходя из данного выше их определения каноническими уравнениями.

Эллипс. Он имеет две взаимно перпендикулярные оси симметрии; их называют его *осями.* Оси пересекаются в его центре симметрии

или — *центре* эллипса. Эти свойства непосредственно видны из того, что эллипс может быть задан уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{1}$$

Здесь осями его служат оси координат, а центром — начало координат.

Действительно, отражение относительно оси x состоит в замене точек $(x,\ y)$ на $(x,\ -y)$, т. е. в перемене знака y. Поэтому, вообще, если y входит в некоторое уравнение только в четных степенях, то кривая, представляемая таким уравнением, симметрична относительно оси x (ср. замечание в конце § 5 гл. I). Аналогично, если уравнение содержит x только в четных степенях, то ось y есть ось симметрии кривой.

Симметрия (или отражение) относительно точки, служащей началом координат, состоит в замене точек $(x,\ y)$ на $(-x,\ -y)$. Поэтому если в уравнение входят члены только с четными степенями или даже четными суммами степеней (т. е. для каждого члена $ax^\alpha y^\beta$ сумма $\alpha+\beta$ четна), то начало является центром симметрии кривой[6].

Точки пересечения эллипса с осями называются его *вершинами*. Величины a, b представляют собой расстояния от центра до вершин эллипса (так как при $y = 0$ из (1) следует, что $x = \pm a$, и при $x = 0$ — что $y = \pm b$). Их называют *полуосями* эллипса. Окружность характеризуется, очевидно, их равенством.

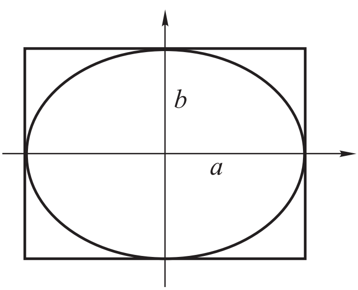


Рис. 27

Ось эллипса, которой отвечает большее из чисел a, b, называется *большой*, другой — *малой*. Обычно предполагают, что в уравнении (1) $a \geqslant b$, так что a — большая полуось, b — малая и ось x — «большая ось», ось y — «малая» (рис. 27).

Из уравнения (1) следует, что $|x| \leqslant a$, $|y| \leqslant b$, т. е. эллипс заключен в прямоугольнике, ограниченном прямыми $x = \pm a$ и $y = \pm b$. Эллипс ограничивает эллиптическую «площадку», которую тоже называют *эллипсом*; она задается неравенством $x^2/a^2 + y^2/b^2 \leqslant 1$. Она выпукла — отрезок, соединяющий любые две ее точки, содержится в ней. (Докажите это.)

[6] Разумеется, все сказанное сохраняет силу, если слово «четный» заменить всюду в этом абзаце на слово «нечетный».

Эллипс как проекция окружности. Эллипсу можно дать следующее определение. Эллипс — это кривая, которая получается при сжатии окружности к какому-либо ее диаметру. При проектировании с плоскости P на плоскость Q происходит сжатие вдоль прямых, перпендикулярных линии пересечения плоскостей P, Q. Поэтому данное определение эллипса равносильно тому, что эллипс есть проекция окружности (рис. 28).

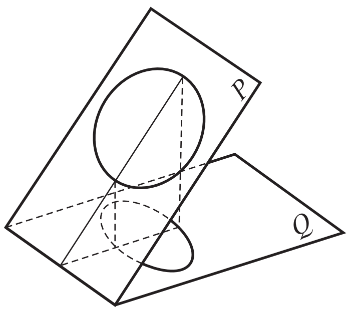


Рис. 28

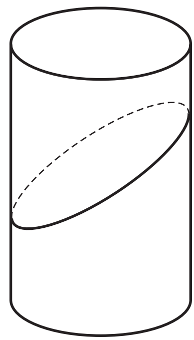


Рис. 29

Сжатие фигуры в k раз к оси x состоит в том, что каждая точка $(x,\ y)$ фигуры переходит в точку $(\widetilde{x},\ \widetilde{y})$ с $\widetilde{x} = x$, $\widetilde{y} = y/k$, так что $x = \widetilde{x}$, $y = k\widetilde{y}$. Поэтому окружность $x^2 + y^2 = a^2$ переходит в кривую

$$\widetilde{x}^2 + k^2\widetilde{y}^2 = a^2,$$

или, если положить $a/k = b$,

$$\frac{\widetilde{x}^2}{a^2} + \frac{\widetilde{y}^2}{b^2} = 1.$$

Совершенно так же растяжение окружности от прямой дает эллипс. Поэтому плоские сечения прямого кругового цилиндра — эллипсы (рис. 29).

Гипербола. Она имеет две оси симметрии и центр симметрии. Это видно из ее уравнения

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{2}$$

Как и для эллипса, осями служат оси координат, центром — начало. Одна ость — «мнимая» — не пересекает гиперболу, другая — «вещественная» — пересекает в двух точках, называемых *вершинами.* (Наглядно это очевидно, а формально доказывается «в два слова». Если

$x = 0$, т. е. берется ось y, то из (2) получается $y^2 = -b^2$, что невозможно. Если $y = 0$, т. е. берется ось x, то $x^2 = a^2$ и $x = \pm a$.)

Следовательно, величина a есть расстояние от центра гиперболы до ее вершин. Она называется *вещественной полуосью*, а b — *мнимой* (потому что при замене b на $\sqrt{-1}b$ уравнение гиперболы переходит в уравнение эллипса). Всегда подразумевается, что a, $b > 0$.

Из уравнения (2) очевидно, что $x^2/a^2 \geqslant 1$. Поэтому $x^2 \geqslant a^2$. Следовательно, на гиперболе нет точек с $|x| < a$. Она состоит из двух ветвей, разделенных целой полосой между прямыми $x = a$, $x = -a$. Эти ветви уходят в бесконечность, так как при $|x| \to \infty$ также $|y| \to \infty$.

Асимптоты гиперболы. Прямые

$$y = \frac{b}{a}x, \quad y = -\frac{b}{a}x \tag{3}$$

обладают в отношении гиперболы с уравнением (2) важными свойствами. Если точка, двигаясь непрерывно по гиперболе, удаляется в бесконечность, то она неограниченно приближается к одной из этих прямых (рис. 30).

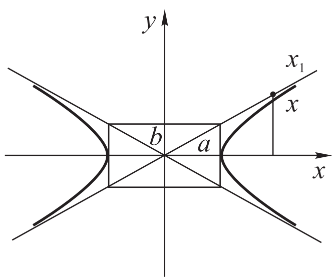


Рис. 30

Прямая, обладающая в отношении какой-либо кривой подобным свойством, называется ее *асимптотой*. Наше утверждение о прямых (3) состоит, стало быть, в том, что они служат асимптотами гиперболы.

Докажем это утверждение. Из уравнения (2) следует, что для точки $X(x,\ y)$, лежащей на гиперболе, имеет место равенство

$$y^2 = \frac{b^2}{a^2}x^2 - b^2.$$

Пусть точка X расположена, например, на той части гиперболы, где $x > 0$, $y > 0$. Вертикальная прямая, проходящая через нее, пересечет прямую $y = \frac{b}{a}x$ в некоторой точке $X_1(x,\ y_1)$, так что

$$y_1^2 = \frac{b^2}{a^2}x^2.$$

Поэтому $y_1^2 - y^2 = b^2$, откуда

$$y_1 - y = \frac{b^2}{y_1 + y}. \qquad (4)$$

Если $x \to +\infty$, то $y \to +\infty$ и $y_1 \to +\infty$; поэтому из равенства (4) следует, что $y_1 - y \to 0$. Тем самым точка X неограниченно приближается к прямой $y = \frac{b}{a}x$.

Остальные части гиперболы, где $x > 0$, $y < 0$ и т. п., не нужно рассматривать, так как из симметрии гиперболы относительно осей ясно, что для них выполнено аналогичное. $\square$

Из доказанного вытекают следующие два

Замечания. 1. Чертить гиперболу надо так, чтобы ее ветви, «распрямляясь», прижимались к асимптотам (рис. 31, *а* — верный, рис. 31, *б* — неверный).

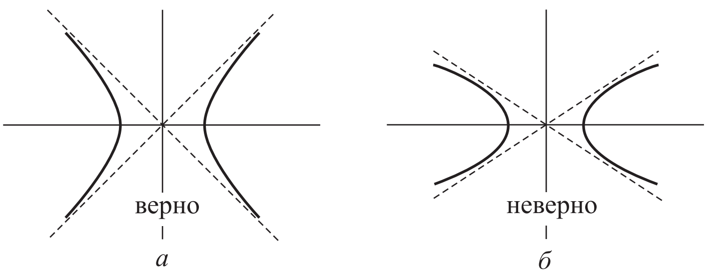


Рис. 31.

2. При движении точки по гиперболе движение это на больших расстояниях все меньше отличается от прямолинейного — от движения по асимптоте.

Парабола. Она имеет одну ось симметрии — это ось x, когда параболу представляет каноническое уравнение $y^2 = 2px$. Точка пересечения параболы с осью называется ее *вершиной.* Каждая прямая, параллельная оси, пересекает параболу в одной точке.

§ 3. Эллипс; его фокальное свойство

Эллипсу можно дать (и очень часто дают) следующее чисто геометрическое определение.

Эллипс — это фигура (кривая), образованная точками, у которых сумма расстояний от двух заданных точек фиксирована (и больше

расстояния между двумя указанными точками). То есть если F_1, F_2 — данные точки, то эллипс образован точками M, для которых

$$F_1M + F_2M = \mathrm{const} > F_1F_2. \qquad (1)$$

Последнее неравенство необходимо: оно означает, что сумма двух сторон треугольника F_1F_2M больше третьей[7].

Точки F_1, F_2 называются *фокусами* эллипса. Если они сливаются, то условие (1) сводится к тому, что $FM = \mathrm{const}$; точки с этим условием образуют окружность. Она считается частным (иногда — вырожденным) случаем эллипса.

Здесь мы будем понимать эллипс в смысле данного определения и докажем, что он задается в подходящих координатах уравнением $x^2/a^2 + y^2/b^2 = 1$ (т. е. является эллипсом в смысле данного в § 1 определения).

Доказательство. Введем прямоугольные координаты с началом посредине между фокусами. т. е. в середине отрезка F_1F_2, и с осью x, проходящей через фокусы (рис. 32). Введем обозначения: постоянную в (1) обозначим $2a$ и положим $F_1M = r_1$, $F_2M = r_2$, так что условие (1) будет выглядеть следующим образом:

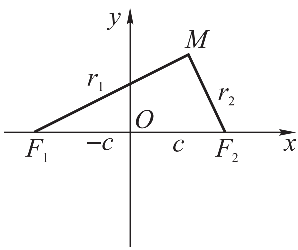


Рис. 32

$$r_1 + r_2 = 2a. \qquad (2)$$

Полагаем еще $F_1F_2 = 2c$, так что расстояние фокусов от начала равно c, и фокусы расположены в точках $F_1(-c,\ 0)$, $F_2(c,\ 0)$. Поэтому

$$r_1^2 = (x+c)^2 + y^2,\ r_2^2 = (x-c)^2 + y^2,\ r_1^2 - r_2^2 = 4cx. \qquad (3)$$

Из (2) находим:

$$r_2^2 = (2a - r_1)^2 = 4a^2 + r_1^2 - 4ar_1,\ 4ar_1 = r_1^2 - r_2^2 + 4a^2. \qquad (4)$$

Отсюда, пользуясь последним равенством в (3), получаем

$$r_1 = \frac{c}{a}x + a. \qquad (5)$$

[7]Точки M с условием $F_1M + F_2M = F_1F_2$ образуют отрезок F_1F_2; точек M с $F_1M + F_2M < F_1F_2$ не существует.

Возводим в квадрат и подставляем r_1^2 из (3). Тогда простые преобразования (которые вы должны проделать!) приводят к уравнению

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1. \tag{6}$$

Так как $a > c$, то можно положить $a^2 - c^2 = b^2$. Тогда уравнение (6) совпадает с уравнением (2) из § 1:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{7}$$

Что мы доказали, выведя это уравнение? Мы доказали, что координаты точек эллипса (в его геометрическом определении условием (2)) удовлетворяют уравнению (7), т. е. что эти точки содержатся среди точек, удовлетворяющих этому уравнению: другими словами — что эллипс содержится в кривой, представленной уравнением (7). Но это еще не значит, что он и есть эта кривая: ведь она может содержать точки, для которых условие (2), определяющее эллипс, не выполняется. Стало быть, нужно еще проверить, что из уравнения (7) вытекает условие (2).

Действуя «в лоб», нужно, пользуясь уравнением (6), вывести выражения для r_1, r_2:

$$r_1 = a + \frac{c}{a}x, \quad r_1 = a - \frac{c}{a}x \tag{8}$$

и, сложив их, получить: $r_1 + r_2 = 2a$. Проделаем этот вывод.

Из выражения (3) для r_1^2 и из уравнения (7) находим:

$$r_1^2 = (x+c)^2 + b^2 - \frac{b^2}{a^2}x^2 = (1 - \frac{b^2}{a^2})x^2 + 2cx + c^2 + b^2,$$

и так как $b^2 = a^2 - c^2$, то

$$r_1^2 = \left(a + \frac{c}{a}x\right)^2. \tag{9}$$

Извлекая корень, получим для r_1 или выражение (8), или то же со знаком минус:

$$r_1 = -\left(a + \frac{c}{a}x\right). \tag{10}$$

Почему (10) не может выполняться? Расстояние $r_1 \geqslant 0$, а правая часть здесь положительна при $a + \frac{c}{a}x < 0$. Однако это невозможно в силу уравнения (7): из него очевидно следует, что $|x| \leqslant a$, но тогда $a + \frac{c}{a}x > 0$, поскольку $0 < c/a < 1$.

Таким образом, извлечение корня из (9) дает для r_1 только выражение (8).

Так как в формулах (3), из которых делается вывод, r_2 отличается от r_1 только знаком при c, то и для r_2 получаем выражение (8). В результате мы получаем, что $r_1 + r_2 = 2a$. Следовательно, уравнение (7) задает эллипс (в данном его геометрическом определении). $\square$

Отметим, что дальше — в § 4 — при аналогичном выводе извлечение корня даст другой результат.

Замечание. Проверку можно провести из геометрических соображений без выкладок. Из уравнения (7), как уже замечено, следует, что $x \in [-a,\ a]$. Каждому x из этого промежутка соответствует определенный y^2 и, стало быть, два значения: $\pm y$. Вместе с тем с каждой стороны от отрезка $|x| \leqslant a$ на перпендикуляре к нему лежит по одной точке M с данной суммой $F_1M + F_2M$ (сумма этих расстояний растет с удалением M от оси x и потому имеет данное значение только в одной точке с каждой стороны от оси). Но, по доказанному при выводе уравнения (7), эти точки лежат на кривой с этим уравнением. А так как на том же перпендикуляре с каждой стороны есть по одной точке этой кривой, то эти точки совпадают с указанными точками. Тем самым эта кривая и есть эллипс, что и требовалось доказать.

Определение. Отношение

$$e = \frac{c}{a} \tag{11}$$

называют *эксцентриситетом* эллипса (потому что оно показывает, насколько фокусы удалены от центра; для окружности $c = 0$, так что $e = 0$).

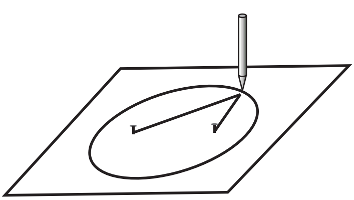


Рис. 33

С этим обозначением формулы (8) записываются компактно:

$$r_1 = a + ex, \quad r_2 = a - ex. \tag{12}$$

Расстояния $r_1 = F_1M$, $r_2 = F_2M$ называют *фокальными радиусами* точки M на эллипсе.

Свойство эллипса, выраженное условием $F_1M + F_2M = \text{const}$, называют его *фокальным свойством.* Поэтому полученный нами результат можно выразить так.

Определения эллипса его фокальным свойством и каноническим уравнением равносильны.

Фокальное свойство дает простой способ чертить эллипс. Воткнув в бумагу две булавки в фокусах эллипса, прикрепляют к ним концами нитку длиной в две большие полуоси. Оттягивая нитку острием карандаша, обводят одну половину эллипса, потом — другую (рис. 33).

§ 4. Гипербола; ее фокальное свойство

Гиперболу можно определить геометрически следующим образом.

Гипербола — это кривая, образованная точками, у которых модуль разности расстояний от двух заданных точек фиксирован (и меньше расстояния между данными точками). То есть если F_1, F_2 — данные точки, то гипербола образуется точками M, для которых

$$|F_1M - F_2M| = \text{const} < F_1F_2. \tag{1}$$

Неравенство здесь выражает, что разность двух сторон треугольника F_1F_2M меньше третьей[8].

Так же, как в случае эллипса, точки F_1, F_2 называются *фокусами* гиперболы, расстояния F_1M, F_2M до точки M на гиперболе называются *фокальными радиусами;* Само свойство гиперболы, выраженное условием (1), называется ее *фокальным свойством.*

Здесь мы будем называть гиперболой кривую, определенную фокальным свойством (1). Задача состоит в том, чтобы доказать, что это определение равносильно доказанному выше в § 1 уравнением $x^2/a^2 - y^2/b^2 = 1$.

Доказательство. Введем обозначения: $F_1M = r_1$, $F_2M = r_2$, $\text{const} = 2a$ и условие (1) запишем в виде $|r_1 - r_2| = 2a$. Это означает, что выполнено одно из двух равенств:

$$r_1 - r_2 = 2a, \quad \text{или} \quad r_2 - r_1 = 2a. \tag{2}$$

Соответственно, гипербола состоит из двух «ветвей»: на одной выполняется первое из этих равенств, на другой — второе.

[8]Точки M, для которых $|F_1M - F_2M| = F_1F_2$, образуют два луча с началом в точках F_1, F_2. Точек M с $|F_1M - F_2M| > F_1F_2$ не существует.

Введем такие же координаты, как в случае эллипса. Начало берем посредине между фокусами, ось x проводим через фокусы. Полагаем $F_1F_2 = 2c$. Фокусы расположены в точках $F_1(-c,\ 0)$, $F_2(c,\ 0)$. Поэтому (так же, как в случае эллипса)

$$r_1^2 = (x+c)^2 + y^2,\ r_2^2 = (x-c)^2 + y^2, \tag{3}$$

$$r_1^2 - r_2^2 = 4cx. \tag{4}$$

Возьмем ту ветвь гиперболы, на которой выполнено первое из равенств (2): $r_1 - r_2 = 2a$. Тогда

$$r_2^2 = (2a - r_1)^2.$$

Отсюда, пользуясь (4), так ж, как в случае эллипса, получаем

$$r_1 = a + \frac{c}{a}x. \tag{5}$$

Возводим в квадрат и подставляем r_1^2 из (3). Тогда простые преобразования приводят, так же как в случае эллипса, к уравнению

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1. \tag{6}$$

Но теперь $c > a$ (вследствие неравенства в условии (1)). Поэтому можно положить $c^2 - a^2 = b^2$, и мы получим:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{7}$$

Мы доказали, что точки, для которых выполняется первое из равенств (2), принадлежат кривой с уравнением (7), т. е. что первая ветвь гиперболы в ней содержится. Но это не вся гипербола. Уравнению (7) удовлетворяет также и вторая ее ветвь. Это вы докажете аналогично предыдущему, исходя из второго равенства (2).

Теперь нужно доказать, что не только обе ветви содержатся в кривой (7), но и что они ее исчерпывают. Тогда и будет доказано, что уравнение (7) действительно задает гиперболу, определенную ее фокальным свойством.

Обратимся к соответствующему выводу, так же как в случае эллипса. Из выражения (3) для r_1^2, пользуясь (7) и тем, что $c^2 = a^2 + b^2$, находим:

$$r_1^2 = (x+c)^2 + \frac{b^2}{a^2}x^2 - b^2 = \left(a + \frac{c}{a}x\right)^2.$$

Отсюда, полагая $c/a = e$ (как и в (11), § 3), получаем

$$r_1 = \pm(a + ex). \tag{8}$$

Минус здесь возможен — в отличие от случая эллипса. Из уравнения (7) очевидно, что $|x| \geqslant a$. Поэтому при $-x \geqslant a$, поскольку $c > a$, будет $-\left(a + \frac{c}{a}x\right) = -a - \frac{c}{a}x \geqslant c - a > 0$.

Аналогично найдем r_2, исходя из выражения (3): для r_2 оно отличается только знаком при c. Поэтому можно сразу написать его, изменяя в (8) знак при e:

$$r_2 = \pm(a - ex). \tag{9}$$

Здесь опять возможны оба знака.

При $x \geqslant a$ имеем

$$r_1 = a + ex, \quad r_2 = -a + ex.$$

При $-x \geqslant a$ имеем

$$r_1 = -(a + ex), \quad r_2 = a - ex.$$

В обоих случаях $|r_1 - r_2| = 2a$.

Итак, мы доказали, что каждая точка кривой с уравнением (7) лежит на одной из ветвей гиперболы. Наше утверждение полностью доказано. □

Таким образом, мы доказали, что ***определения гиперболы ее фокальным свойством и каноническим уравнением равносильны.***

Фокальное свойство гиперболы дает простой способ чертить ее. Втыкаем в бумагу две булавки в фокусах гиперболы и прикрепляем к ним — к одной и другой — две нитки; на них надеваем колечко и, натянув нитки, помещаем его в точке между фокусами, где должна быть вершина гиперболы. Если теперь перемещать колечко, натягивая нитки, то каждая из них — от булавки до колечка — будет удлиняться на одну и ту же величину, так что разность их будет постоянна. Поэтому острие карандаша, вставленного в колечко или прижатого к нему, будет описывать ветвь гиперболы (рис. 34).

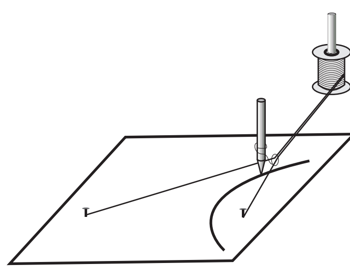


Рис. 34

§ 5. Парабола; ее фокус и директриса. Директрисы эллипса и гиперболы

Параболе можно дать следующее геометрическое определение.

Парабола — это кривая, образованная точками, равноудаленными от данной точки и от данной прямой (при условии, что данная точка не лежит на этой прямой). То есть если F — данная точка, D — данная прямая, а r и d — расстояния от точки M до F и D, то парабола образована точками M, для которых

$$r = d. \tag{1}$$

Точка F называется *фокусом* параболы, прямая D — ее *директрисой.*

Покажем, что парабола определяется указанным геометрическим свойством.

Доказательство. Введем координаты, проведя ось x перпендикулярно директрисе через фокус. Начало возьмем в середине отрезка между директрисой и фокусом. Длину этого отрезка (расстояние от фокуса F до директрисы) обозначим p, так что фокус будет точкой $(p/2,\ 0)$, а директриса — прямой $x = -p/2$ (рис. 35). Расстояние от точки $M(x,\ y)$ до директрисы будет

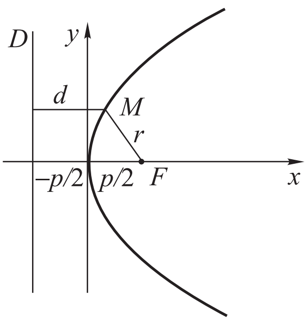


Рис. 35

$$d = \left|x + \frac{p}{2}\right|. \tag{2}$$

Расстояние же $r = FM$ до фокуса будет

$$r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2}. \tag{3}$$

Из равенства $r = d$ получаем $d^2 = r^2$, т. е.

$$x^2 + px + \frac{p^2}{4} = x^2 - px + \frac{p^2}{4} + y^2, \tag{4}$$

откуда

$$y^2 = 2px. \tag{5}$$

Этим доказано, что точки с $r = d$ содержатся в кривой с уравнением $y^2 = 2px$. Обратное доказывается «обратным ходом»: из (5) получаем (4), и, извлекая корни, получаем $r = d$. $\square$

Итак, определения параболы равенством $r = d$ и уравнением (5) равносильны.

Директрисы эллипса и гиперболы. Эллипс, отличный от окружности, и гипербола характеризуются вместе с параболой общим свойством. Это кривые, образуемые точками, у которых отношение расстояний до данной точки F и до данной прямой D (не проходящей через F) одно и то же. То есть если r и d — расстояния точки M до F и D, то

$$\frac{r}{d} = e = \text{const}. \tag{6}$$

Причем для эллипса $e < 1$, а для гиперболы $e > 1$, и это число есть не что иное, как эксцентриситет. Для параболы же, как следует из предыдущего, $e = 1$.

Прямая D также называется *директрисой*, а точка F — *фокусом* кривой — эллипса или гиперболы. По симметрии эллипса и гиперболы у них по два фокуса и две директрисы. Это те самые фокусы, которые фигурируют в фокальном свойстве. Директриса и фокус, для которых выполнено (6), располагаются с одной стороны от центра. Директрисы перпендикулярны линии фокусов и расположены от центра на расстоянии

$$q = \frac{a^2}{c} = \frac{a}{e}.$$

Кривая с постоянным отношением (6), очевидно, определяется двумя величинами: отношением e в (6) (эксцентриситет) и расстоянием p фокуса от директрисы. Для эллипса и гиперболы

$$p = \frac{b^2}{c}. \tag{7}$$

Окружность исключается: для нее $c = 0$; у нее нет директрисы.

Докажем, что эллипс, отличный от окружности, и гипербола обладают указанным свойством. Рассмотрим рис. 36. Введем координаты x, y как и раньше так, чтобы фокусы были в точках $F_1(-c,\ 0)$, $F_2(c,\ 0)$. Директриса D_2 — это прямая $x = q = a/e$. У точек эллипса $|x| \leqslant a$ и $e < 1$. Поэтому расстояние произвольной его точки $M(x,\ y)$ от директрисы D_2 будет

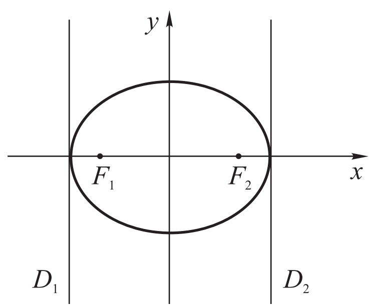


Рис. 36

$$d_2 = q - x = \frac{a}{e} - x. \tag{8}$$

Вместе с тем по формуле (12) § 3 фокальное расстояние

$$r_2 = a - ex. \tag{9}$$

Поэтому $r_2 = ed_2$, что и требовалось доказать.

По симметрии эллипса точно так же $r_1 = ed_1$. Стало быть, равенство (6) на эллипсе выполняется.

Обратно: из (6) и (8) следует (9), а из (9), как показано выше в § 3, вытекает уравнение эллипса. Таким образом, равенство (6) с $e < 1$ характеризует эллипс, что и требовалось доказать.

Аналогичный вывод для гиперболы читатель проведет сам. Здесь нужно иметь в виду, что для одного фокуса есть две формулы расстояния (по формулам (8), (9) § 4). Для каждого из этих «расстояний» нужно проверить равенство $r/d = e$. □

§ 6. Уравнение в полярных координатах

Возьмем полярные координаты с началом в фокусе рассматриваемой кривой — эллипса, гиперболы или параболы, а начальный луч направим через ближайшую к фокусу вершину или, что равносильно, в сторону соответствующей директрисы перпендикулярно к ней (рис. 37). При таком выборе координат радиальная координата точки на кривой есть не что иное, как фокальный радиус.

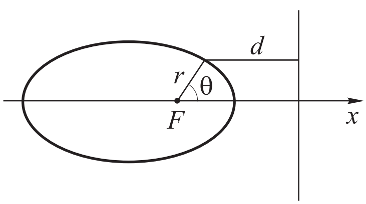


Рис. 37

По свойству директрисы и фокуса

$$r = ed. \tag{1}$$

Пусть q — расстояние от фокуса до директрисы, а x — координата точки, когда ось x направлена по начальному лучу, т. е. к директрисе. Тогда если $x < q$, то

$$d = q - x. \tag{2}$$

Так будет для эллипса, параболы и той ветви гиперболы, которая ближе к взятому фокусу. Вместе с тем в полярных координатах

$$x = r \cos\theta.$$

Поэтому (1) дает

$$r = ed = e(q - x) = eq - er \cos\theta,$$

откуда

$$r = \frac{eq}{1 + e\,\cos\theta}. \qquad (3)$$

Это и есть ***уравнение*** эллипса, параболы и одной ветви гиперболы ***в полярных координатах.***

Величина eq имеет простой геометрический смысл. При $\theta = \pi/2$ формула (3) дает $r = eq$; это значит, что eq есть расстояние от фокуса до кривой по перпендикуляру к оси (рис. 38). Оно называется *фокальным параметром* и обозначается p.

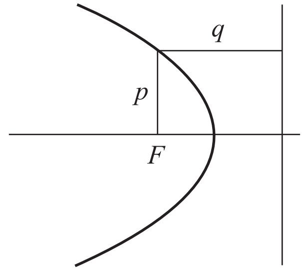


Рис. 38

В случае гиперболы для другой ее ветви (рис. 39)

$$d = x - q.$$

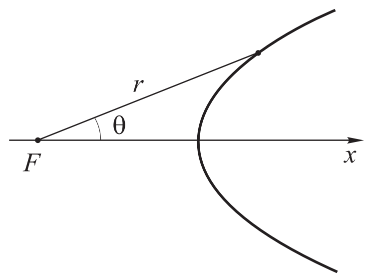


Рис. 39

Поэтому

$$r = ed = e(x - q) = er\,\cos\theta - eq,$$

откуда

$$r = \frac{-eq}{1 - e\,\cos\theta} = \frac{eq}{e\,\cos\theta - 1}. \qquad (4)$$

Для параболы $e = 1$ и $q = p$, это то самое p, которое входит в ее уравнение.

Введем фокальный параметр p в уравнения (3), (4):

$$r = \frac{p}{1 + e\,\cos\theta}, \quad r = \frac{p}{e\,\cos\theta - 1}. \qquad (5)$$

Однако чаще начальный луч полярных координат направляют противоположно, от ближайшей вершины. Тогда угол θ изменяется на π, и косинус меняет знак. Поэтому при таких координатах уравнения будут

$$r = \frac{p}{1 - e\,\cos\theta}, \quad r = \frac{-p}{1 + e\,\cos\theta}. \qquad (6)$$

При задании уравнениями (5) начальный луч $\theta = 0$ проходит через точки, ближайшие к фокусу. При задании уравнениями (6), напротив, начальный луч не пересекает ни параболу, ни гиперболу. Гиперболу не пересекают лучи, для которых знаменатель неположителен. (В каком угле заключены эти лучи?)

Значение уравнения конических сечений в полярных координатах. Согласно первому закону Кеплера (открытому им около 1610 г.), планеты двигаются по эллипсам, в фокусе которых находится Солнце.

Потом Ньютон вывел из своего «второго закона» и закона тяготения, что всякое тело под влиянием притяжения Солнца необходимо двигается по кривой — эллипсу, параболе или ветви гиперболы — с фокусом в Солнце, которая охватывает фокус (как на рис. 37), не считая того случая, когда оно движется прямо к Солнцу или от него.

Естественно принять за начало координат центр Солнца и в качестве одной из координат взять расстояние от него. В качестве другой естественно взять угол поворота радиус-вектора, идущего из центра Солнца к движущемуся телу. Тем самым мы вводим полярные координаты в плоскости, в которой расположена траектория небесного тела. Так естественно появляется представление этой траектории в полярных координатах. Так же естественно, что уравнение (3) задает именно ветвь гиперболы, обходящую фокус: траектория не может состоять из двух разделенных ветвей и из-за притяжения тела к Солнцу, находящемуся в фокусе, должна огибать его.

Вообще уравнение (3) представляет любую возможную траекторию материальной точки во всяком поле, где сила притяжения направлена к фиксированному центру и обратно пропорциональна квадрату расстояния (исключая случай, когда точка движется по прямой к центру или от него). Если же действует сила отталкивания от центра, — тоже обратно пропорциональная квадрату расстояния, — то точка под влиянием такой силы движется по другой ветви гиперболы, отходящей от фокуса (рис. 39).

По закону «обратного квадрата расстояния» действуют «кулоновские» силы: силы притяжения между противоположными по знаку электрическими зарядами и отталкивания — между одноименными зарядами.

Еще в начале XX века великий английский физик Резерфорд из наблюдений над рассеянием альфа-частиц заключил, что атом имеет ядро, которое положительно заряжено. В этом важнейшем откры-

тии существенную роль сыграло то, что траектории α-частиц должны быть ветвями гипербол[9].

Таким образом, конические сечения и представление их уравнениями (6) связаны с самыми фундаментальными проблемами естествознания: с познанием законов движения небесных тел, с одной стороны, и открытием ядра атома, с другой!

Уравнение конического сечения, отнесенное к вершине. Если начало координат взять в вершине и сохранить направление оси x, получим уравнение, общее для эллипса, гиперболы и параболы:

$$y^2 = 2px + (e^2 - 1)x^2.$$

Здесь p, e — как и выше, в уравнениях (5), (6), — фокальный параметр и эксцентриситет. При $e < 1$ получается эллипс, при $e > 1$ — гипербола, при $e = 1$ — парабола[10]. Выведите это уравнение и рассмотрите, как изменяется вид кривой при постоянном p и непрерывном изменении e, начиная от нуля.

§ 7. Классификация КВП

В § 1 мы указали 8 типов КВП. Теперь докажем, что *ими исчерпываются все возможные КВП*, т. е. что уравнение

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_1x + 2a_2y + a_0 = 0 \qquad (1)$$

может представлять только кривую одного из указанных типов. Это можно назвать «основной теоремой о КВП».

Идея доказательства состоит в том, чтобы путем преобразований координат привести уравнение (1) к возможно более простому виду, когда тип кривой, какую оно может представлять, станет очевидным. Действительно, уравнение может быть сложным потому, что система

[9] Конечно, наблюдаемые траектории — прямые: ветвь гиперболы далеко от вершины неотличима от асимптоты. Но движение по гиперболе определяет направление движения летящей к ядру и отброшенной ядром α-частицы. Это и играет роль в выводе формулы Резерфорда, применяя которую к анализу экспериментальных данных, он доказал существование ядра.

[10] При $e < 1$ $(e^2 - 1)x^2 < 0$, при $e > 1$ $(e^2 - 1)x^2 > 0$. Это выражают сами греческие названия: «эллипс» — недостаток, «гипербола» — избыток. Название «парабола» означает «прикладывание» и связано с геометрическим смыслом уравнения $y^2 = 2px$: квадрат, построенный на отрезке y, равновелик прямоугольнику со сторонами $2p$, x.

координат не связана естественно с кривой. Задача состоит в том, чтобы перейти от, так сказать, «случайной» системы координат к «естественной» для данной кривой.

Лемма (А). *Всегда можно повернуть оси координат так, чтобы член с произведением xy исчез.*

Доказательство. Допустим, $a_{12} \neq 0$. Введем координаты x', y', повернув оси на угол α. Тогда

$$x = x' \cos\alpha - y' \sin\alpha,$$
$$y = x' \sin\alpha + y' \cos\alpha.$$

Подставляя эти выражения в (1), найдем коэффициент $2a'_{12}$ при произведении $x'y'$; он будет

$$2a'_{12} = 2a_{22} \sin\alpha \cos\alpha - 2a_{11} \sin\alpha \cos\alpha + \\ + 2a_{12}(\cos^2\alpha - \sin^2\alpha) = (a_{22} - a_{11}) \sin 2\alpha + 2a_{12} \cos 2\alpha.$$

Так как предположено $a_{12} \neq 0$, то, взяв α так, что

$$\frac{\cos 2\alpha}{\sin 2\alpha} = \operatorname{ctg} 2\alpha = \frac{a_{11} - a_{22}}{2a_{12}},$$

получим $2a'_{12} = 0$. $\square$

Таким образом, мы можем предполагать дальше, что уравнение имеет вид

$$a_{11}x^2 + a_{22}y^2 + 2a_1x + 2a_2y + a_0 = 0. \qquad (2)$$

Понятно, что в этом уравнении коэффициенты, вообще говоря, не такие, как в (1), и координаты x, y берутся уже в другой системе, но для простоты записи мы этого не отмечаем.

Лемма (В). *Если в уравнении* (2) $a_{11} \neq 0$ ($a_{22} \neq 0$), *то член с* $x\,(y)$ *можно исключить переносом начала координат.*

Доказательство. Перенесем начало в точку $(-a_1/a_{11},\ 0)$, так что $x = x' - a_1/a_{11}$, $y = y'$.

Тогда

$$a_{11}x^2 + 2a_1x = \\ = a_{11}(x')^2 - 2a_{11}\frac{a_1}{a_{11}}x' + \left(\frac{a_1}{a_{11}}\right)^2 + 2a_1x' = \\ = a_{11}(x')^2 + \left(\frac{a_1}{a_{11}}\right)^2,$$

т. е. член с x' отсутствует. (Полезно отметить, что при переносе начала, т. е. подстановке $x = x' + p$ (и $y = y' + q$), коэффициенты при x^2, y^2 уже не изменяются.) □

Теперь можно заметить, что в уравнении (2) хотя бы один из коэффициентов a_{11}, a_{22} не равен нулю. Можно считать, что $a_{22} \neq 0$ (иначе переименуем координаты). Тогда, применяя «операцию В», т. е. указанный только что в лемме (В) перенос начала, мы придем к уравнению, где член с y' отсутствует. Таким образом, оказывается, что всегда можно выбрать координаты так, чтобы данная КВП представлялась уравнением вида

$$a_{11}x^2 + a_{22}y^2 + 2a_1x + a_0 = 0. \tag{3}$$

Из этого уравнения видно, что ось x служит осью симметрии кривой. Стало быть, мы доказали, между прочим, что всякая КВП имеет хотя бы одну ось симметрии!

Теперь рассмотрим различные возможности для уравнения (3).

Случай I. Пусть в (3) $a_{11} \neq 0$ (и $a_{22} \neq 0$, как уже предположено). Тогда операцией (В) можно прийти к уравнению без члена с x, так что уравнение (3) примет вид

$$a_{11}x^2 + a_{22}y^2 + a_0 = 0 \tag{4}$$

(новые координаты мы для простоты записи обозначаем x, y).

Теперь есть две возможности: (I, 1) $a_0 \neq 0$, (I, 2) $a_0 = 0$.

(I, 1) Допустим, $a_0 \neq 0$. Тогда (4) можно переписать в виде

$$ax^2 + \beta y^2 = 1 \quad \left(\alpha = -\frac{a_{11}}{a_0},\ \beta = -\frac{a_{22}}{a_0}\right). \tag{5}$$

Тут есть следующие возможности:

(I, 1а) Пусть $\alpha > 0$ и $\beta > 0$. Тогда, полагая $\alpha = a^{-2}$, $\beta = b^{-2}$, приведем (5) к виду

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

т. е. кривая — эллипс.

(I, 1б) Пусть $\alpha < 0$, $\beta < 0$. Тогда равенство (5) невозможно ни при каких x, y, так что кривая — пустая.

(I, 1в) Пусть α, β разных знаков. Пусть именно $\alpha > 0$, $\beta < 0$. Тогда, полагая $\alpha = a^{-2}$, $\beta = -b^{-2}$, получим

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

т. е. кривая — гипербола. (Понятно, что если $\alpha < 0$, $\beta > 0$, то можно переименовать координаты.)

(I, 2) Пусть $a_0 = 0$. Тогда (4) сводится к

$$a_{11}x^2 + a_{22}y^2 = 0, \tag{6}$$

и есть две возможности:

(I, 2а) Пусть a_{11}, a_{22} разных знаков, так что $a_{11}/a_{22} < 0$ (по условию $a_{11} \neq 0$, $a_{22} \neq 0$). Тогда, деля (6) на a_{22} и полагая $-a_{11}/a_{22} = k^2$, заменяем (6) на

$$y^2 - k^2x^2 = 0, \quad \text{т. е.} \quad (y + kx)(y - kx) = 0.$$

Таким образом, кривая есть пара прямых $y = kx$, $y = -kx$.

(I, 2б) Пусть a_{11}, a_{22} одного знака. Тогда, как очевидно, (6) возможно лишь при $x = y = 0$, т. е. кривая есть точка.

Случай II. В уравнении (3) $a_{11} = 0$ ($a_{22} \neq 0$ по условию).

Поэтому уравнение (3) (если разделить на a_{22}) сводится к уравнению вида

$$y^2 + 2px + q = 0. \tag{7}$$

Мы различаем прежде всего два случая: либо $p \neq 0$, либо $p = 0$.

(II, 1) Пусть $p \neq 0$. Тогда перенесем начало в точку $(-q/2p,\ 0)$, т. е. положим $x = x' - q/2p$. В результате уравнение (7) приведется к виду

$$y^2 + 2px' = 0,$$

или, после снятия штриха, получаем

$$y^2 = -2px,$$

т. е. кривая — парабола. Если $p > 0$, то, обращая направление оси x, т. е. изменяя знак x, получим справа $2px$.

(II, 2) Пусть $p = 0$. Тогда (7) приводится к виду

$$y^2 + q = 0 \tag{8}$$

и имеется три случая: $q < 0$, $q = 0$, $q > 0$.

(II, 2а) Если $q < 0$, то (8) равносильно

$$(y + \sqrt{-q})(y - \sqrt{-q}) = 0,$$

т. е. кривая есть пара параллельных прямых $y = \sqrt{-q}$, $y = -\sqrt{-q}$.

(II, 2б) Если $q = 0$, то уравнение (8) сводится к $y^2 = 0$, т. е. $y = 0$, т. е. представляет прямую.

(II, 2в) Если $q > 0$, то равенство (8) невозможно ни при каком y, т. е. кривая — пустая. Но пустую кривую мы уже получили раньше в случае (I, 1б).

«Основная теорема о КВП» доказана. □

Замечание. Пустая кривая получилась у нас дважды: в случаях (I, 1б), (II, 2в). Это объясняется тем, что ее представляют в этих случаях существенно различные уравнения. В случае (I, 1б) ее уравнение содержит обе координаты, а в случае (II, 2в) — только одну.

Отметим, что уравнение второй степени может представлять окружность только тогда, когда оно имеет вид

$$a(x^2 + y^2) + 2a_1x + 2a_2y + a_0 = 0. \tag{9}$$

Действительно, если кривая — окружность, то после поворота осей, исключающего xy, должно оказаться $a_{11} = a_{22}$. (Эллипс с равными осями.) Но $x^2 + y^2$ не изменяется при повороте осей, т. е. $x^2 + y^2 = \widetilde{x}^2 + \widetilde{y}^2$. Поэтому и до поворота должно быть $a_{11} = a_{22}$, $a_{12} = 0$, т. е. уравнение должно иметь указанный вид.

Однако уравнение (9) может представлять окружность лишь при $a_1^2 + a_2^2 > aa_0$; ее радиус $r = \sqrt{a_1^2 + a_2^2 - aa_0}$. (Докажите это.) При $a_1^2 + a_2^2 = aa_0$ уравнение задает точку, а при $a_1^2 + a_2^2 < aa_0$ — пустое множество.

Отсюда, между прочим, следует, что если дано уравнение (9) и заранее известно, что представляемая им кривая содержит хотя бы две точки, то она — окружность.

Аналогично, если в общем уравнении (1) $a_{12} = 0$ и a_{11}, a_{22} одного знака и кривая содержит хотя бы две точки, то она — эллипс. (Почему?)

Глава III

ВЕКТОРЫ И КООРДИНАТЫ

§ 1. Понятие вектора

Направленные отрезки. Термин «вектор» употребляют в геометрии по крайней мере в двух смыслах, — хотя и связанных, но тем не менее существенно различных. С одной стороны, вектором называют направленный отрезок, с другой стороны, вектор понимают, как

понимают в физике «векторные величины» — скорость, силу и т. п. (когда говорят, например, что одна и та же сила может быть приложена к разным телам — в разных точках). Представляется удобным различать соответственно «конкретный вектор» — направленный отрезок — и «абстрактный (или, как принято говорить, свободный) вектор». Рассмотрим сначала направленные отрезки — конкретные векторы.

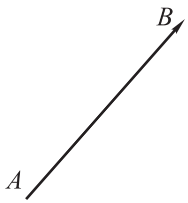


Рис. 40

Направленным отрезком называют отрезок, у которого один конец именуют *началом*, а другой — *концом* этого направленного отрезка. Поэтому в шутку можно сказать, что направленный отрезок — это отрезок, у которого только один конец, а другой переименован в начало. Рисуют направленный отрезок со стрелкой на конце и обозначают $\overrightarrow{AB}$, если A — его начало, а B — конец (рис. 40).

Замечание. Данное определение может показаться не математическим. Принято говорить, что направленным отрезком называется отрезок, у которого указан порядок концов: первый конец считаем началом, второй — концом направленного отрезка. Однако что значит «порядок концов», кроме того, что одному концу отнесено слово «первый», другому — «второй»? Поэтому относить одному концу слово «начало» — ничуть не менее строго, и, кстати, это соответствует тому, что говорят «начало вектора», но никто не говорит «первый конец вектора».

Таким образом, данное нами определение направленного отрезка вполне математическое, оно только необычно с точки зрения принятой «научности» выражений.

Говорят: «направленный отрезок лежит на прямой» или «направленный отрезок перпендикулярен (параллелен) прямой» (или плоскости или другому направленному отрезку и т. п.), если соответствующий ему отрезок (т. е., в сущности, он сам) содержится в прямой, перпендикулярен (параллелен) прямой (или плоскости и т. п.). Точно так же под *длиной* направленного отрезка понимают длину соответствующего ему отрезка.

Направленный отрезок называют, как уже было сказано, *вектором*, и мы будем так поступать, хотя бы ради краткости. Кроме обозначения началом и концом: $\overrightarrow{AB}$, $\overrightarrow{PQ}$ и т. п., векторы обозначают $\overrightarrow{a}$, $\overrightarrow{b}$ и т. п., или полужирным шрифтом: $\boldsymbol{a}$, $\boldsymbol{b}$ и т. п.

Все даваемые здесь определения и выводы относятся к векторам в пространстве, равно как и на плоскости.

Направление и равенство направленных отрезков. Векторы — направленные отрезки — $\overrightarrow{AB}$, $\overrightarrow{CD}$ называются *одинаково направленными* или, короче, *сонаправленными*, если при неограниченном продолжении они не расходятся, т. е., точно говоря, каковы бы ни были равные отрезки AM и CN, содержащие точки B и D, расстояние MN не превосходит некоторого данного расстояния (рис. 41). Это соответствует практическому представлению о «движении параллельными курсами» (не сонаправленные векторы AM, CN могут сначала сходиться, но потом станут расходиться, рис. 42). Сонаправленные векторы обозначаются: $\overrightarrow{AB} \uparrow\uparrow \overrightarrow{CD}$ и т. п. Параллельные, но не сонаправленные векторы называются *противоположно направленными* и обозначаются: $\overrightarrow{AB} \uparrow\downarrow \overrightarrow{CD}$.

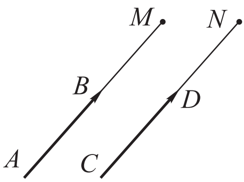


Рис. 41

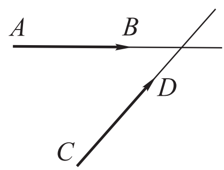


Рис. 42

Теорема 1. *Два вектора, сонаправленные с третьим, сонаправлены друг с другом.*

Доказательство. Пусть векторы $\overrightarrow{AB}$, $\overrightarrow{CD}$ сонаправлены с $\overrightarrow{EF}$. Тогда существуют такие числа d_1, d_2, что если отрезки AM, CN, EP равны и содержат, соответственно, точки B, D, F, то

$$MP < d_1, \quad NP < d_2.$$

Но по неравенству треугольника

$$MP + NP \geqslant MN.$$

Поэтому

$$MN < d_1 + d_2,$$

и это верно для любых точек M, N таких, что $B \in AM$, $D \in CN$,

$AM = CN$[11]. Следовательно, векторы $\overrightarrow{AB}$, $\overrightarrow{CD}$ сонаправлены, что и требовалось доказать. □

Определение. Два вектора считаются *равными*, если они сонаправлены и имеют одну и ту же длину. Равенство векторов записывается, как принято для равенства: $\overrightarrow{AB} = \overrightarrow{CD}$ и т. п.

Теорема 2. *Два вектора, равные третьему, равны друг другу.*

Доказательство. Это непосредственно следует из определения равенства благодаря теореме 1. Действительно, пусть $\overrightarrow{AB} = \overrightarrow{CD}$ и $\overrightarrow{EF} = \overrightarrow{CD}$. Это значит, во-первых, что $\overrightarrow{AB}$ и $\overrightarrow{EF}$ сонаправлены с $\overrightarrow{CD}$ и, следовательно, по теореме 1 сонаправлены друг с другом. Во-вторых, длины $\overrightarrow{AB}$ и $\overrightarrow{EF}$ равны длине $\overrightarrow{CD}$, а значит, и сами равны. Итак, $\overrightarrow{AB}$ и $\overrightarrow{EF}$ сонаправлены и имеют равные длины, а значит, они, по определению, равны. □

Следствие. *Отношение равенства векторов рефлексивно, симметрично и транзитивно, т. е.*: 1) $\boldsymbol{a} = \boldsymbol{a}$; 2) *если* $\boldsymbol{a} = \boldsymbol{b}$, *то* $\boldsymbol{b} = \boldsymbol{a}$; 3) *если* $\boldsymbol{a} = \boldsymbol{b}$ *и* $\boldsymbol{b} = \boldsymbol{c}$, *то* $\boldsymbol{a} = \boldsymbol{c}$.

Первые два свойства заключены в определении, третье вытекает из теоремы 2. □

Отношение, обладающее этими свойствами, называют вообще *эквивалентностью* — отношением эквивалентности. Для векторов мы называем его *равенством.*

Замечание. Существует мнение, что говорить о равенстве векторов в этом смысле нельзя. Потому что в других случаях «равные» означают «совпадающие». Например, $x = 2$ означает, что x есть число два. Но $\overrightarrow{AB} = \overrightarrow{CD}$ вовсе не означает, что $\overrightarrow{AB}$ совпадает с $\overrightarrow{CD}$. Поэтому в некоторых книгах векторы, «равные» в нашем смысле, называют «эквиполлентными». Однако мы вместе с многими другими авторами считаем вполне возможным применить к векторам термин «равенство», как оно было здесь определено. То, что равенство для векторов (направленных отрезков) и для чисел имеет разный смысл, не должно вас смущать, нужно только понимать это различие.

О сонаправленных и равных векторах.

Теорема 3. *Два вектора* $\overrightarrow{AB}$, $\overrightarrow{CD}$, *не лежащие на одной прямой, сонаправлены тогда и только тогда, когда они параллельны и лежат в содержащей их плоскости по одну сторону от прямой* AC, *проходящей через их начала. Если же векторы* $\overrightarrow{AB}$, $\overrightarrow{CD}$ *лежат на одной*

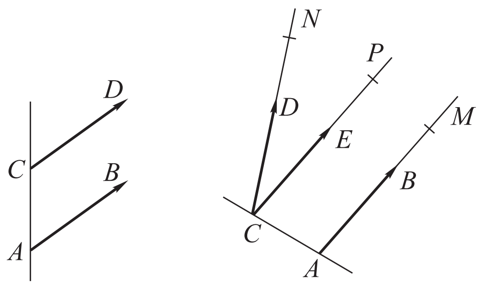
---

[11] Неравенство треугольника, $AB + BC \geqslant AC$, доказывают в планиметрии, но оно верно и в пространстве; через любые три точки A, B, C проходит плоскость, и потому для них верно все, что доказывается в планиметрии.

прямой, то они сонаправлены тогда и только тогда, когда один из лучей AB, CD содержит другой.

Доказательство. Второе утверждение теоремы можно считать очевидным. Обратимся к первому (рис. 43, *а*).

Пусть отрезки AB, CD параллельны и лежат по одну сторону от прямой AC. Тогда если вдоль них проведены равные отрезки AM, CN, то мы получаем параллелограмм со стороной MN, противоположной AC, так что $MN = AC$. Стало быть, векторы $\overrightarrow{AB}$, $\overrightarrow{CD}$ сонаправлены.



а *б*

Рис. 43

Докажем обратное. Пусть векторы $\overrightarrow{AB}$, $\overrightarrow{CD}$ не лежат на одной прямой и сонаправлены. Проведем из точки C отрезок CE, параллельный AB и лежащий по ту же сторону от прямой AC. По доказанному $\overrightarrow{CE} \uparrow\uparrow \overrightarrow{AB}$.

Допустим, CE не налегает на CD, т. е. отрезки CE, CD образуют угол. Тогда если вдоль них откладывать все бо́льшие равные отрезки CP, CN, то расстояние PN будет неограниченно расти, а если равные им отрезки AM откладывать вдоль AB, то все время $MP = AC$ (рис. 43, *б*).

По неравенству треугольника

$$MN > NP - MP = NP - AC,$$

и так как NP не ограничено, то и MN не ограничено. Но это противоречит тому, что по условию $\overrightarrow{AB}$ и $\overrightarrow{CD}$ сонаправлены, так что MN должно быть ограничено. Таким образом, отрезок CE должен налегать на CD. А стало быть, CD расположен так же, как CE, т. е. тоже параллелен AB и лежит по ту же сторону от AC. Что и требовалось доказать. $\square$

Теорема 4. *Для всякого вектора $\overrightarrow{AB}$ при любой точке C существует вектор $\overrightarrow{CD}$, равный $\overrightarrow{AB}$, и притом только один. Короче это выражают так: от всякой точки можно отложить вектор, равный данному, и притом только один.*

Доказательство. Пусть даны вектор $\overrightarrow{AB}$ и точка C, не лежащая на прямой AB. Проведем через точки A, B, C плоскость, в ней проведем отрезок CD, параллельный и равный AB и расположенный по ту же сторону от прямой AC, что и отрезок AB (рис. $43, a$). По теореме 3, $\overrightarrow{CD} = \overrightarrow{AB}$. Вектор, равный $\overrightarrow{AB}$, отложен от точки C. Он единственный, потому что через точку C проходит только одна прямая, параллельная данной AB. Вектор $\overrightarrow{CD}$ должен лежать на ней, и его направление и длина определены однозначно.

Если точка C лежит на прямой AB, то можно поступить, например, так. Проводим отрезок AC. Если он совпадает по направлению с $\overrightarrow{AB}$, то откладываем $CD = AB$ вдоль CA. В противном случае откладываем $CD = AB$ в противоположную сторону. То, что здесь направленный отрезок $\overrightarrow{CD}$, равный $\overrightarrow{AB}$, только один, очевидно: на луче от его начала откладывается только один отрезок, равный данному. □

Отметим важное свойство равенства векторов.

Лемма. *Равенство $\overrightarrow{AB} = CD$ имеет место тогда и только тогда, когда $\overrightarrow{AC} = \overrightarrow{BD}$.*

Доказательство. Если отрезки AB, CD не лежат на одной прямой, то согласно теореме 3 равенство $\overrightarrow{AB} = \overrightarrow{CD}$ равносильно тому, что отрезки AB, CD служат противоположными сторонами параллелограмма. Две другие его стороны — это отрезки AC и BD, так что $\overrightarrow{AC} = \overrightarrow{BD}$ (рис. 44). Совершенно так же из равенства $\overrightarrow{AC} = \overrightarrow{BD}$ следует, что $\overrightarrow{AB} = \overrightarrow{CD}$.

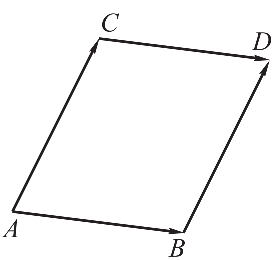


Рис. 44

Если точки A, B, C, D лежат на одной прямой, то доказательство получается, если посмотреть, как отрезки AC, BD получаются из AB, CD. Рассмотрение разных случаев читатель проведет сам. □

Нуль-вектор. В доказанной только что лемме есть логический пропуск. Равенство $\overrightarrow{AB} = \overrightarrow{CD}$ может оказаться равенством $\overrightarrow{AB} = \overrightarrow{AB}$, так как не оговорено, что точка C отлична от A. Равенство $\overrightarrow{AB} = \overrightarrow{AB}$

верно, а по лемме (поскольку $C = A$ и $D = B$) оно имеет место тогда и только тогда, когда $\overrightarrow{AA} = \overrightarrow{BB}$.

Поэтому для того, чтобы лемма была верна во всех случаях, нужно присоединить к векторам (направленным отрезкам) «нуль-векторы», у которых начало и конец совпадают. Все они считаются равными друг другу. Каждая точка представляет собой нуль-вектор, если ее рассматривать как его начало и конец. Длина нуль-вектора равна нулю, а направления он не имеет (!).

Вообще, вместо векторов — направленных отрезков можно рассматривать «векторы» — *упорядоченные пары точек:* одна точка — «начало», другая — «*конец*», не исключая их совпадения. В некоторых случаях так векторы и определяют (как пары точек), и это бывает более удобным. Теоретически это безразлично, так как направленный отрезок определяется своими началом и концом. Достоинство направленного отрезка перед парой точек — в наглядности.

Свободный вектор.

Определение. *Свободным вектором* (или просто *вектором*) называется абстрактный объект, связанный с равными направленными отрезками тем, что каждый из равных друг другу направленных отрезков считается представителем данного свободного вектора, а неравные направленные отрезки представляют разные свободные векторы.

Так понимаемый вектор называется свободным потому, что он представляется направленным отрезком независимо от того, от какой точки тот отложен. Равные направленные отрезки $\overrightarrow{AB}$, $\overrightarrow{A'B'}$, отложенные от разных точек, представляют один и тот же вектор. Так что говорят: $\overrightarrow{AB}$ — это вектор $\boldsymbol{a}$, отложенный от точки A, а $\overrightarrow{A'B'}$ — тот же вектор, отложенный от точки A'. Это можно сравнить, например, с силой, когда она рассматривается независимо от того, к какому телу она приложена; как, скажем, сила тяги локомотива; ее можно приложить к разным поездам. Так и вектор можно «прилагать» — откладывать от разных точек.

В частности, все нуль-векторы $\overrightarrow{AA}$, $\overrightarrow{BB}$ и т. п. представляют один и тот же свободный нуль-вектор, который обозначается $\mathbf{0}$.

Данное определение вектора в виде абстрактного объекта, представляемого направленными отрезками, выражает то, как на самом деле понимают вектор в геометрии. Дальше это будет особенно ясно видно при определении сложения векторов. Векторы — это объекты, с которыми производят определенные действия, прежде всего

сложение. Учитывая это, мы дальше уточним определение вектора.

Точное наглядное представление вектора дается перемещением материальной точки или, более грубо, материального тела. Перемещение определяется направлением и расстоянием, на какое оно происходит: на 2 км к северу, 10 шагов вправо и т. п. Данное конкретное перемещение материальной точки из геометрической точки A в B изображается направленным отрезком $\overrightarrow{AB}$. Но одно и то же перемещение могут совершать разные тела из разных мест, и их перемещения изобразятся равными направленными отрезками. Так, люди в строю, делающие по команде два шага вперед, совершают одно и то же перемещение: «два шага вперед».

Вектор, как и перемещение, *характеризуется направлением и расстоянием или длиной. Задать вектор — значит задать направление и расстояние (длину).* Направление задается направленным отрезком или лучом, и все направленные отрезки, сонаправленные с данным, имеют данное направление. Поскольку у равных направленных отрезков одно и то же направление и одна и та же длина, то они представляют один и тот же вектор. Фраза «Вектор $\boldsymbol{a}$ отложен от точки A» значит, что построен направленный отрезок $\overrightarrow{AB}$, представляющий этот вектор, и мы будем писать $\overrightarrow{AB} = \boldsymbol{a}$. *Длина* или *модуль* вектора — это длина представляющего его направленного отрезка, она обозначается как модуль $|\boldsymbol{a}|$ и т. п.

Замечание. Данное здесь определение вектора как абстрактного объекта, представляемого направленными отрезками, может показаться необычным. Однако все общие понятия имеют такой характер. Что такое «дом» вообще как не абстрактный объект, связанный с конкретными домами? И как бывают разные типы домов, так бывают и разные «типы» векторов — неравные абстрактные векторы.

Однако такой взгляд в математике непривычен, вместо этого пользуются понятием множества и дают определение: вектор — это множество всех равных направленных отрезков, и даже хотят представить его множеством равных направленных отрезков, отложенных от всех точек пространства (хотя едва ли кто-нибудь на самом деле так представляет вектор). В духе этой теоретико-множественной установки нужно определять так: «дом» — это множество всех домов: число два — это множество всех пар предметов, и т. п. На самом же деле вектор — это представляемый направленными отрезками элемент векторной алгебры.

§ 2. Сложение векторов

Определение. Если материальная точка переместилась из точки A в точку B, а потом из B в C, то получается перемещение из A в C. Поэтому, естественно, считается, что направленные отрезки $\overrightarrow{AB}$, $\overrightarrow{BC}$, представляющие эти перемещения, складываются и дают в сумме направленный отрезок $\overrightarrow{AC}$ (рис. 45). Это записывается так:

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

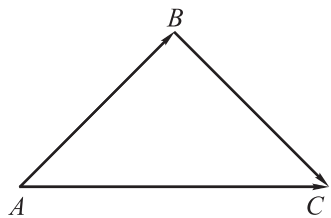


Рис. 45

В соответствии с этим определяют сложение (свободных) векторов. Пусть даны два вектора $\boldsymbol{a}$, $\boldsymbol{b}$. Откладываем $\boldsymbol{a}$ от какой-либо точки A и получаем направленный отрезок $\overrightarrow{AB} = \boldsymbol{a}$. Затем от точки B откладываем вектор $\boldsymbol{b}$ и получаем $\overrightarrow{BC} = \boldsymbol{b}$. Тогда направленный отрезок $\overrightarrow{AC}$ представляет вектор $\boldsymbol{c}$, который и принимается за сумму векторов $\boldsymbol{a}$ и $\boldsymbol{b}$: $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}$.

Однако это определение связано с произвольно выбранной точкой A. Поэтому нужно еще доказать, что сумма — вектор $\boldsymbol{a}+\boldsymbol{b}$ — не зависит от выбора точки A. То есть нужно доказать

Корректность. *Независимо от того, с какой точки A начинается построение суммы, оно дает равные направленные отрезки $\overrightarrow{AC}$ и тем самым — один и тот же вектор* $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}$. Короче,

если $\overrightarrow{AB} = \overrightarrow{A_1B_1}$ *и* $\overrightarrow{BC} = \overrightarrow{B_1C_1}$, *то* $\overrightarrow{AC} = \overrightarrow{A_1C_1}$.

Доказательство. По лемме, доказанной в предыдущем параграфе, равенства $\overrightarrow{AB} = \overrightarrow{A_1B_1}$ и $\overrightarrow{BC} = \overrightarrow{B_1C_1}$ равносильны равенствам $\overrightarrow{AA_1} = \overrightarrow{BB_1}$, $\overrightarrow{BB_1} = \overrightarrow{CC_1}$.

Отсюда, по транзитивности равенства, $\overrightarrow{AA_1} = \overrightarrow{CC_1}$. А отсюда по той же лемме, на которую только что сослались, следует $\overrightarrow{AC} = \overrightarrow{A_1C_1}$, что и требовалось доказать. □

Описанная операция получения вектора $\boldsymbol{c}$ по векторам $\boldsymbol{a}$, $\boldsymbol{b}$ называется их ***сложением по правилу треугольника.*** При всей наглядности это правило страдает тем недостатком, что в нем сложение производится в определенном порядке, так что еще нужно доказать, что сумма не зависит от порядка слагаемых. Но это непосредственно

видно при ***сложении по правилу параллелограмма.*** Оно производится так.

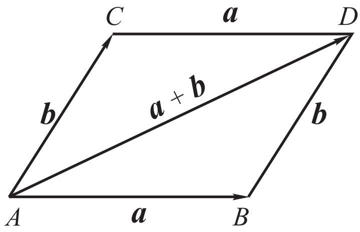


Рис. 46

Если векторы $\boldsymbol{a}$, $\boldsymbol{b}$ не параллельны, их откладывают от одной точки, получая $\overrightarrow{AB} = \boldsymbol{a}$, $\overrightarrow{AC} = \boldsymbol{b}$, и на отрезках AB, AC строят параллелограмм. Его диагональ AD и дает сумму $\boldsymbol{a}+\boldsymbol{b}$ (рис. 46). Действительно, в параллелограмме $\boldsymbol{a} = \overrightarrow{AB} = \overrightarrow{CD}$, $\boldsymbol{b} = \overrightarrow{AC} = \overrightarrow{BD}$. Поэтому

$$\boldsymbol{c} = \overrightarrow{AD} = \overrightarrow{AB} + \overrightarrow{BD} = \boldsymbol{a} + \boldsymbol{b},$$

и точно так же

$$\boldsymbol{c} = \overrightarrow{AD} = \overrightarrow{AC} + \overrightarrow{CD} = \boldsymbol{b} + \boldsymbol{a}.$$

Стало быть, $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$.

Если векторы откладываются на одной прямой, то равенство $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b}+\boldsymbol{a}$ выводится из сложения длин, если $\boldsymbol{a} \uparrow\uparrow \boldsymbol{b}$, и из их вычитания, если $\boldsymbol{a} \uparrow\downarrow \boldsymbol{b}$. Читатель сам проделает этот вывод.

Сложение по правилу треугольника соответствует сложению перемещений, совершаемых одно вслед за другим. Сложение по правилу параллелограмма соответствует сложению двух одновременных перемещений, как, например, при движении по палубе плывущего корабля. В физике пользуются правилом параллелограмма, поскольку складываются векторы, например силы, приложенные в одной точке.

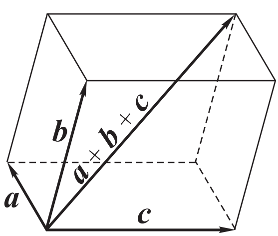


Рис. 47

Подобно правилу параллелограмма, отметим *правило параллелепипеда:* сумма трех векторов, не параллельных одной плоскости, представляется диагональю параллелепипеда, построенного на этих векторах, отложенных от одной точки, как ребрах (рис. 47).

Замечание. Если не вводить понятие свободного вектора, а определять вектор только как направленный отрезок, то сложение векторов, вообще говоря, невозможно, так как складываются не данные направленные отрезки, а только равные им. Складывать можно направленные отрезки с общим началом.

Алгебраические свойства сложения векторов.

1. Для каждых двух векторов $\boldsymbol{a}$, $\boldsymbol{b}$ однозначно определена их сумма $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}$.

2. Сложение векторов переместительно (коммутативно), т. е. $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$. (Это доказано.)

3. Сложение векторов сочетательно (ассоциативно), т. е. $(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c})$.

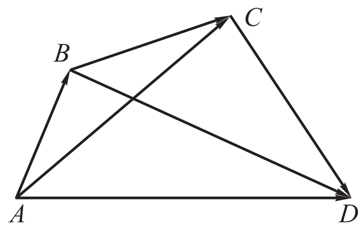


Рис. 48

Доказательство очевидно, если складывать по правилу треугольника (рис. 48):

$$(\overrightarrow{AB} + \overrightarrow{BC}) + \overrightarrow{CD} = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD} = \\ = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AB} + (\overrightarrow{BC} + \overrightarrow{CD}).$$

4. Прибавление нуль-вектора ничего не меняет:

$$\boldsymbol{a} + \mathbf{0} = \boldsymbol{a},$$

так как

$$\overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB}.$$

5. Для каждого вектора есть противоположный $-\boldsymbol{a}$, т. е. такой, что

$$\boldsymbol{a} + (-\boldsymbol{a}) = \mathbf{0}.$$

(Если $\boldsymbol{a} = \overrightarrow{AB}$, то $-\boldsymbol{a} = \overrightarrow{BA}$ и $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA}$.)

Все эти свойства означают, что векторы образуют *коммутативную группу* относительно сложения (то, что векторная операция записывается как сложение, указывают, говоря, что это группа «в аддитивной записи»).

6. Однозначно определено вычитание, т. е. для каждых двух векторов $\boldsymbol{a}$, $\boldsymbol{b}$ существует такой вектор $\boldsymbol{c}$ — их *разность*, что $\boldsymbol{a} = \boldsymbol{b} + \boldsymbol{c}$; пишут $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$.

Действительно, если сложить $\boldsymbol{c} = \boldsymbol{a} + (-\boldsymbol{b})$, то по определению это и будет разность векторов $\boldsymbol{a}$ и $\boldsymbol{b}$, так как

$$\boldsymbol{b} + \boldsymbol{c} = \boldsymbol{b} + \boldsymbol{a} + (-\boldsymbol{b}) = \boldsymbol{a} + (\boldsymbol{b} + (-\boldsymbol{b})) = \boldsymbol{a} + \mathbf{0} = \boldsymbol{a}.$$

Другой разности нет, так как если $\boldsymbol{a} = \boldsymbol{b} + \boldsymbol{c}$, то, прибавляя $(-\boldsymbol{b})$ к обеим сторонам равенства, получим

$$\boldsymbol{a} + (-\boldsymbol{b}) = \boldsymbol{b} + \boldsymbol{c} + (-\boldsymbol{b}) = \boldsymbol{c} + \boldsymbol{b} + (-\boldsymbol{b}) = \boldsymbol{c} + \mathbf{0} = \boldsymbol{c}.$$

Разложение вектора на составляющие. Самолет пошел на посадку. Его перемещение состоит из двух составляющих: вертикальной и горизонтальной: первая показывает, насколько он снизился, вторая — насколько и в каком направлении он переместился над землей за время снижения (рис. 49, *а*).

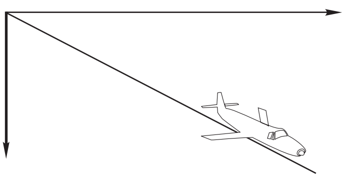


Рис. 49*а*

Вес груза, висящего на треноге, разлагается вдоль ног треноги на три составляющие (рис. 49, *б*). Вес тела на наклонной плоскости разлагается на давление на плоскость и «скатывающую силу» вдоль плоскости (рис. 49, б).

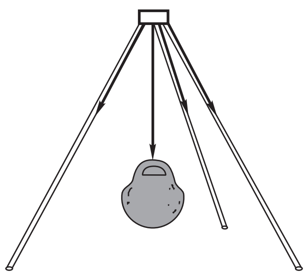


Рис. 49*б*

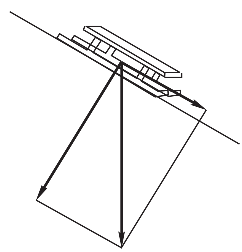


Рис. 49*в*

Коротко говоря, *составляющие* данного вектора — это такие векторы, сумма которых равна этому вектору. Он «составляется» из них, как сумма из слагаемых, и разлагается на них, как на слагаемые. Поэтому и говорят о разложении на составляющие.

Разложение вектора в плоскости. *Всякий вектор в плоскости однозначно разлагается на составляющие, параллельные двум данным пересекающимся прямым.*

Доказательство. Пусть даны прямые a, b, пересекающиеся в точке O, и в той же плоскости — вектор $\boldsymbol{v}$. Отложим его от точки O, так что будет $\overrightarrow{OV} = \boldsymbol{v}$. Пусть $\overrightarrow{OV}$ не лежит ни на одной из прямых a,

b, проведем из точки V отрезки $VA||b$, $VB||a$ с концами A на a и B на b (рис. 50). Тогда по правилу параллелограмма

$$\overrightarrow{OV} = \overrightarrow{OA} + \overrightarrow{OB}.$$

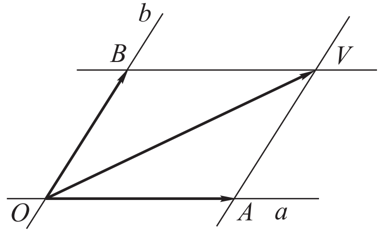


Рис. 50

Векторы $\boldsymbol{v}_a = \overrightarrow{OA}$, $\boldsymbol{v}_b = \overrightarrow{OB}$ и являются составляющими вектора $\boldsymbol{v}$ вдоль прямых a, b: $\boldsymbol{v}_a||a$, $\boldsymbol{v}_b||b$, $\boldsymbol{v}_a + \boldsymbol{v}_b = \boldsymbol{v}$.

Если $\boldsymbol{v}||a$, так что вектор $\boldsymbol{v} = \overrightarrow{OV}$ оказывается на прямой a, то он сам и служит своей составляющей $\boldsymbol{v}_a$ вдоль a, а $\boldsymbol{v}_b = \mathbf{0}$, так что $\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_b$. Если же $\boldsymbol{v}||b$, то $\boldsymbol{v}_a = \mathbf{0}$ и $\boldsymbol{v}_b = \boldsymbol{v}$.

Итак, доказано, что вектор разлагается на составляющие по прямым a, b. Нужно доказать, что такое разложение единственно.

Пусть имеются два разложения, так что

$$\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_b = \boldsymbol{v}'_a + \boldsymbol{v}'_b. \tag{1}$$

Отсюда $\boldsymbol{v}_a - \boldsymbol{v}'_a = \boldsymbol{v}'_b - \boldsymbol{v}_b$.

Здесь вектор $\boldsymbol{v}_a - \boldsymbol{v}'_a$ параллелен прямой a, а $\boldsymbol{v}'_b - \boldsymbol{v}_b$ параллелен прямой b. Ясно, что они могут быть равны, только если они нулевые, т. е. если $\boldsymbol{v}_a - \boldsymbol{v}'_b = \mathbf{0}$, $\boldsymbol{v}'_b - \boldsymbol{v}_b = \mathbf{0}$, так что $\boldsymbol{v}'_a = \boldsymbol{v}_a$, $\boldsymbol{v}'_b = \boldsymbol{v}_b$. Значит оба разложения (1) совпадают и двух разных быть не может, что и требовалось доказать. □

Разложение вектора по прямой и плоскости. *Всякий вектор* $\boldsymbol{v}$ *допускает, и притом единственное, разложение на две составляющие: одну —* $\boldsymbol{v}_a$, *параллельную данной прямой* a, *другую —* $\boldsymbol{v}_\alpha$, *параллельную данной плоскости* α, *пересекающей прямую* a; *в формулах:*

$$\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_\alpha, \quad \boldsymbol{v}_a||a, \ \boldsymbol{v}_\alpha||\alpha.$$

Доказательство. Пусть $\boldsymbol{v}$ — данный вектор, a и α — данные прямая и плоскость, пересекающиеся в точке O. Отложив вектор $\boldsymbol{v}$ от точки O, получим $\overrightarrow{OV} = \boldsymbol{v}$ (рис. 51).

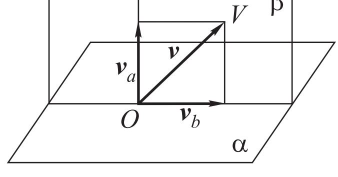


Рис. 51

Пусть β — плоскость, проходящая через прямые a и OV. Она пересекает α по некоторой прямой b. По доказанному вектор $\boldsymbol{v} = \overrightarrow{OV}$ разлагается

на составляющие по прямым a и b: $\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_b$. Но $\boldsymbol{v}_b$ есть, очевидно, его составляющая на плоскости α: $\boldsymbol{v}_b = \boldsymbol{v}_\alpha$. Следовательно,

$$\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_\alpha.$$

Единственность такого разложения доказывается аналогично предыдущему. □

Разложение вектора по трем прямым. *Всякий вектор* $\boldsymbol{v}$ *допускает, и притом единственное, разложение на составляющие, параллельные трем данным прямым* a, b, c, *не параллельным одной плоскости:*

$$\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_b + \boldsymbol{v}_c, \quad \boldsymbol{v}_a || a,\ \boldsymbol{v}_b || b,\ \boldsymbol{v}_c || c.$$

Доказательство. Пусть даны три прямые, не параллельные одной плоскости. Перенесем их, если нужно так, чтобы они пересеклись в одной точке O; обозначим их a, b, c (рис. 52). Пусть α — плоскость, проходящая через прямые b, c. Данный вектор разложим на составляющие $\boldsymbol{v}_a$, $\boldsymbol{v}_\alpha$ по прямой a и плоскости α. Затем разложим $\boldsymbol{v}_\alpha$ по прямым b, c. В результате получим разложение $\boldsymbol{v}$ на составляющие:

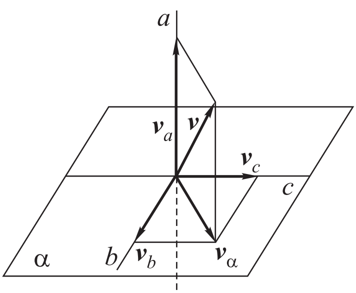


Рис. 52

$$\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_b + \boldsymbol{v}_c,$$
$$\boldsymbol{v}_a || a,\ \boldsymbol{v}_b || b,\ \boldsymbol{v}_c || c.$$

Такое разложение единственно, так как по доказанному выше разложение $\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_\alpha$ единственно и разложение $\boldsymbol{v}_\alpha = \boldsymbol{v}_b + \boldsymbol{v}_c$ единственно. □

Общий вывод о разложении на составляющие. Полученные выводы можно суммировать:

Теорема 1. *Всякий вектор допускает, и притом единственное, разложение на составляющие в каждом из трех следующих случаев:*

а) *по двум пересекающимся прямым, если вектор параллелен проходящей через них плоскости;*

б) *по пересекающимся прямой и плоскости;*

в) *по трем прямым, не параллельным одной плоскости.* □

Из этой теоремы вытекает

Теорема 2. *При сложении векторов соответствующие составляющие складываются.*

Доказательство. Докажем эту теорему, например, для разложения векторов по прямой a и пересекающей ее плоскости α. Пусть даны векторы $\boldsymbol{u}$, $\boldsymbol{v}$, и пусть $\mathbf{w} = \boldsymbol{u} + \boldsymbol{v}$. Разложив эти векторы по прямой a и плоскости α, имеем

$$\boldsymbol{u} = \boldsymbol{u}_a + \boldsymbol{u}_\alpha, \quad \boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_\alpha, \quad \boldsymbol{w} = \boldsymbol{w}_a + \boldsymbol{w}_\alpha.$$

Тогда из первых двух разложений

$$\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v} = \boldsymbol{u}_a + \boldsymbol{u}_\alpha + \boldsymbol{v}_a + \boldsymbol{v}_\alpha = (\boldsymbol{u}_a + \boldsymbol{v}_a) + (\boldsymbol{u}_\alpha + \boldsymbol{v}_\alpha).$$

Здесь вектор $\boldsymbol{u}_a + \boldsymbol{v}_a$ параллелен прямой a, а $\boldsymbol{u}_\alpha + \boldsymbol{v}_\alpha$ — параллелен плоскости α, т. е. они представляют разложение вектора $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{w}$. Поэтому, ввиду единственности разложения, $\boldsymbol{w}_a = \boldsymbol{u}_a + \boldsymbol{v}_a$, $\boldsymbol{w}_\alpha = \boldsymbol{u}_\alpha + \boldsymbol{v}_\alpha$, что и требовалось доказать.

В случае других разложений доказательство совершенно такое же. □

Доказанную теорему можно перефразировать так:

Векторы можно складывать по составляющим.

§ 3. Умножение вектора на число. Координаты вектора

Коллинеарность и компланарность. Свободные векторы называют *коллинеарными*, если представляющие их направленные отрезки параллельны некоторой прямой. При этом удобно считать нуль-вектор коллинеарным любому свободному вектору. Обозначение — то же, что и для параллельности: $\boldsymbol{a} || \boldsymbol{b}$ и т. д.

Свободные векторы называют *компланарными*, если представляющие их направленные отрезки параллельны некоторой плоскости. При этом, опять же, удобно считать нуль-вектор параллельным любой плоскости. Поэтому три вектора, среди которых есть нуль-вектор, всегда компланарны.

Другими словами, если векторы откладывать из одной точки, то коллинеарные векторы окажутся лежащими на одной прямой, а компланарные — в одной плоскости.

Умножение вектора на число. *Произведением* вектора $\boldsymbol{a}$ на действительное число x называется вектор, обозначенный $x\boldsymbol{a}$ или $\boldsymbol{a}x$, такой, что, во-первых, его длина равна $|x||\boldsymbol{a}|$ и, во-вторых, если $\boldsymbol{a} \neq \mathbf{0}$, то при $x > 0$ $x\boldsymbol{a} \uparrow\uparrow \boldsymbol{a}$ и при $x < 0$ $x\boldsymbol{a} \uparrow\downarrow \boldsymbol{a}$ (если $\boldsymbol{a} = \mathbf{0}$ или $x = 0$, то, по первому условию, $|x\boldsymbol{a}| = 0$, так что $x\boldsymbol{a} = \mathbf{0}$).

Теорема 1. *Ненулевой вектор $\boldsymbol{a}$ и вектор $\boldsymbol{b}$ коллинеарны тогда и только тогда, когда существует такое число x, что $\boldsymbol{b} = x\boldsymbol{a}$; такое число x при данных $\boldsymbol{a}$ и $\boldsymbol{b}$ единственное.*

Доказательство. Если $\boldsymbol{a} \neq \boldsymbol{0}$ и $\boldsymbol{b} = x\boldsymbol{a}$, то $\boldsymbol{b}||\boldsymbol{a}$, как следует из самого определения умножения вектора на число (по которому либо $\boldsymbol{b} = \boldsymbol{0}$, либо $\boldsymbol{b} \uparrow\uparrow \boldsymbol{a}$, либо $\boldsymbol{b} \uparrow\downarrow \boldsymbol{a}$).

Пусть $\boldsymbol{a} \neq \boldsymbol{0}$ и $\boldsymbol{b}||\boldsymbol{a}$. Найдем такое x, что $\boldsymbol{b} = x\boldsymbol{a}$. Если $\boldsymbol{b} = \boldsymbol{0}$, то $x = 0$. Если $\boldsymbol{b} \uparrow\uparrow \boldsymbol{a}$, то берем $x = |\boldsymbol{b}|/|\boldsymbol{a}|$. Рассмотрим вектор $x\boldsymbol{a} = \frac{|\boldsymbol{b}|}{|\boldsymbol{a}|}\boldsymbol{a}$. По определению умножения на число

$$|x\boldsymbol{a}| = \frac{|\boldsymbol{b}|}{|\boldsymbol{a}|}|\boldsymbol{a}| = |\boldsymbol{b}|,$$

и так как $x > 0$, то $x\boldsymbol{a} \uparrow\uparrow \boldsymbol{a}$. А так как $\boldsymbol{b} \uparrow\uparrow \boldsymbol{a}$, то $x\boldsymbol{a} \uparrow\uparrow \boldsymbol{b}$. Таким образом, $|x\boldsymbol{a}| = |\boldsymbol{b}|$ и $x\boldsymbol{a} \uparrow\uparrow \boldsymbol{b}$, следовательно, $x\boldsymbol{a} = \boldsymbol{b}$.

Если $\boldsymbol{b} \uparrow\downarrow \boldsymbol{a}$, то берем $x = -|\boldsymbol{b}|/|\boldsymbol{a}|$, и тогда из определения умножения на число

$$|x\boldsymbol{a}| = \frac{|\boldsymbol{b}|}{|\boldsymbol{a}|}|\boldsymbol{a}| = |\boldsymbol{b}|, \text{ и так как } x\boldsymbol{a} \uparrow\downarrow \boldsymbol{a}, \text{ то } x\boldsymbol{a} \uparrow\uparrow \boldsymbol{b}.$$

Следовательно, $x\boldsymbol{a} = \boldsymbol{b}$. Теорема доказана. □

Координаты вектора.

Теорема 2. *Пусть в плоскости даны два неколлинеарных вектора $\boldsymbol{a}$, $\boldsymbol{b}$. Тогда всякий вектор $\boldsymbol{x}$ в этой плоскости может быть представлен в виде*

$$\boldsymbol{x} = x\boldsymbol{a} + y\boldsymbol{b},$$

и притом единственным образом.

(Говорят: вектор $\boldsymbol{x}$ разлагается по векторам $\boldsymbol{a}$, $\boldsymbol{b}$.)

Доказательство. Отложим векторы $\boldsymbol{a}$, $\boldsymbol{b}$ от одной точки O: получим $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$, и проведем прямые OA, OB — обозначим их a, b. Согласно теореме 2 предыдущего параграфа всякий вектор в той же плоскости однозначно разлагается на составляющие, параллельные прямым a, b:

$$\boldsymbol{x} = \boldsymbol{x}_a + \boldsymbol{x}_b. \tag{1}$$

Так как векторы $\boldsymbol{a}$, $\boldsymbol{b}$ параллельны прямым a, b, то по доказанной только что теореме 1 векторы $\boldsymbol{x}_a$, $\boldsymbol{x}_b$ выражаются через векторы $\boldsymbol{a}$, $\boldsymbol{b}$:

$$\boldsymbol{x}_a = x\boldsymbol{a}, \quad \boldsymbol{x}_b = y\boldsymbol{b}. \tag{2}$$

Таким образом, из (1) следует

$$\boldsymbol{x} = x\boldsymbol{a} + y\boldsymbol{b}. \tag{3}$$

Это представление единственно, поскольку представления (1), (2) единственны. □

Теорема 3. *Если векторы* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ *некомпланарны, то всякий вектор* $\boldsymbol{x}$ *выражается через них в виде*

$$\boldsymbol{x} = x\boldsymbol{a} + y\boldsymbol{b} + z\boldsymbol{c}.$$

Доказательство. Отложив векторы $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ от одной точки O, получим $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$, $\overrightarrow{OC} = \boldsymbol{c}$ и проведем прямые OA, OB, OC; обозначим их a, b, c. По теореме 2 предыдущего параграфа любой вектор $\boldsymbol{x}$ однозначно разлагается на составляющие по этим прямым:

$$\boldsymbol{x} = \boldsymbol{x}_a + \boldsymbol{x}_b + \boldsymbol{x}_c. \tag{4}$$

По теореме 1 существуют такие числа x, y, z, что

$$\boldsymbol{x}_a = x\boldsymbol{a}, \quad \boldsymbol{x}_b = y\boldsymbol{b}, \quad \boldsymbol{x}_c = z\boldsymbol{c}, \tag{5}$$

причем такие числа единственны. Из (4) и (5)

$$\boldsymbol{x} = x\boldsymbol{a} + y\boldsymbol{b} + z\boldsymbol{c}, \tag{6}$$

причем x, y, z определены однозначно, так как представления (5), (4) единственны. □

Определение. Числа x, y в теореме 2 и x, y, z в теореме 3 называются *координатами* вектора $\boldsymbol{x}$ относительно векторов $\boldsymbol{a}$, $\boldsymbol{b}$ (в *базисе* $\boldsymbol{a}$, $\boldsymbol{b}$) и соответственно относительно векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ (в *базисе* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$).

Теорема 2 утверждает: на плоскости при неколлинеарных векторах $\boldsymbol{a}$, $\boldsymbol{b}$ всякий вектор $\boldsymbol{x}$ имеет относительно них определенные координаты; он выражается через них по формуле (3). Очевидно, что вместе с тем по той же формуле всякие два числа, заданные в фиксированном порядке, определяют некоторый вектор $\boldsymbol{x}$.

Аналогично, теорема 3 утверждает: при некомпланарных векторах $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ в пространстве всякий вектор $\boldsymbol{x}$ имеет относительно них определенные координаты; он выражается через них по формуле (6). Очевидно, вместе с тем, что всякая упорядоченная тройка чисел x, y, z определяет по той же формуле некоторый вектор $\boldsymbol{x}$.

Определение. Векторы $\boldsymbol{a}$, $\boldsymbol{b}$ на плоскости и $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ — в пространстве называются *основными* (или, как говорят, они образуют *базис*).

Значение координат вектора состоит не только в том, что они задают вектор, но и в том, что операции сложения и умножения вектора на число представляются такими же операциями с координатами, т. е. справедлива следующая

Теорема 4. *При сложении векторов их координаты складываются, так что координаты суммы равны суммам соответственных координат.*

При умножении вектора на число его координаты умножаются на то же число, так что координаты произведения вектора на число равны произведениям координат на это число.

Для того чтобы доказать эту теорему, нужно установить основные свойства умножения вектора на число.

Алгебраические свойства умножения вектора на число. Из определения умножения вектора на число непосредственно вытекают три свойства умножения.

1. Если $\boldsymbol{a} = \boldsymbol{0}$ или $x = 0$, то $x\boldsymbol{a} = \boldsymbol{0}$ (так как по первому условию $|x\boldsymbol{a}| = |x||\boldsymbol{a}|$).

2. $1 \cdot \boldsymbol{a} = \boldsymbol{a}$ для всякого вектора $\boldsymbol{a}$.

3. $(-1) \cdot \boldsymbol{a} = \boldsymbol{a}$.

Далее, умножение вектора на число обладает тремя «вычислительными» свойствами или, как еще говорят, для него выполняются следующие три «закона».

4. Сочетательный закон (ассоциативность): для любого вектора $\boldsymbol{a}$ и любых чисел x, y

$$x(y\boldsymbol{a}) = (xy)\boldsymbol{a}.$$

5. Распределительный закон (дистрибутивность) для численных множителей: для любого вектора $\boldsymbol{a}$ при любых числах x, y

$$(x + y)\boldsymbol{a} = x\boldsymbol{a} + y\boldsymbol{a}.$$

6. Распределительный закон (дистрибутивность) для умножаемых векторов: для любого числа x и любых векторов $\boldsymbol{a}$, $\boldsymbol{b}$

$$x(\boldsymbol{a} + \boldsymbol{b}) = x\boldsymbol{a} + x\boldsymbol{b}.$$

Доказательство. (4) Достаточно рассмотреть случай, когда $x \neq 0$, $y \neq 0$, $\boldsymbol{a} \neq \boldsymbol{0}$. Сравним векторы $(xy)\boldsymbol{a}$ и $x(y\boldsymbol{a})$. Длины их очевидно равны. Если $xy > 0$, то либо $x > 0$, $y > 0$, либо $x < 0$, $y < 0$.

В первом случае умножение на y и x не изменяет направления, во втором — изменяет его дважды, так что по-прежнему $x(y\boldsymbol{a}) \uparrow\uparrow (xy)\boldsymbol{a}$.

Если $xy < 0$, то один из сомножителей, скажем, $x > 0$, а другой $y < 0$. Поэтому при умножении на y, потом на x направление изменяется так, что $x(y\boldsymbol{a}) \uparrow\downarrow \boldsymbol{a}$ как и $(xy)\boldsymbol{a}$.

Таким образом, во всех случаях $x(y\boldsymbol{a}) \uparrow\uparrow (xy)\boldsymbol{a}$. $\square$

Докажем первый распределительный закон (5):

$$(x+y)\boldsymbol{a} = x\boldsymbol{a} + y\boldsymbol{a}.$$

Доказательство. Если одно из чисел x, y равно нулю или $\boldsymbol{a} = \boldsymbol{0}$, то формула очевидна. Поэтому допустим, что $x \neq 0$, $y \neq 0$, $\boldsymbol{a} \neq \boldsymbol{0}$. Если x, y одного знака, то векторы $x\boldsymbol{a}$, $y\boldsymbol{a}$, $(x+y)\boldsymbol{a}$ одинаково направлены и длина вектора $x\boldsymbol{a} + y\boldsymbol{a}$ равна сумме длин векторов $x\boldsymbol{a}$ и $y\boldsymbol{a}$, как и длина вектора $(x+y)\boldsymbol{a}$. Это легко вычислить. Так как x, y одного знака, то $|x+y| = |x| + |y|$. Поэтому

$$|(x+y)\boldsymbol{a}| = |x+y||\boldsymbol{a}| = (|x|+|y|)|\boldsymbol{a}| = |x||\boldsymbol{a}| + |y||\boldsymbol{a}|.$$

С другой стороны, так как x, y одного знака, то векторы $x\boldsymbol{a}$, $y\boldsymbol{a}$ одинаково направлены, и потому длины их складываются, так что

$$|x\boldsymbol{a} + y\boldsymbol{a}| = |x\boldsymbol{a}| + |y\boldsymbol{a}| = |x||\boldsymbol{a}| + |y||\boldsymbol{a}|.$$

Итак, векторы $(x+y)\boldsymbol{a}$ и $x\boldsymbol{a} + y\boldsymbol{a}$ сонаправленны, и длины их равны, поэтому они сами равны, что и требовалось доказать.

Допустим теперь, что x, y разных знаков. Если при этом $x + y = 0$, то

$$(x+y)\boldsymbol{a} = \boldsymbol{0} \quad \text{и} \quad x\boldsymbol{a} + y\boldsymbol{a} = x\boldsymbol{a} - x\boldsymbol{a} = \boldsymbol{0}.$$

Если же $x + y \neq 0$, то либо $(-x)$, либо $(-y)$ того же знака, что $x + y$. Допустим, это $(-y)$. Тогда по доказанному

$$(x+y)\boldsymbol{a} - y\boldsymbol{a} = (x+y-y)\boldsymbol{a} = x\boldsymbol{a},$$

откуда $(x+y)\boldsymbol{a} = x\boldsymbol{a} + y\boldsymbol{a}$. $\square$

Докажем второй распределительный закон (6):

$$x(\boldsymbol{a} + \boldsymbol{b}) = x\boldsymbol{a} + x\boldsymbol{b}.$$

Доказательство. Если $x = 0$ или хотя бы один из векторов $\boldsymbol{a}$, $\boldsymbol{b}$ нулевой, то равенство очевидно. Поэтому допустим, что $x \neq 0$, $\boldsymbol{a} \neq \boldsymbol{0}$,

$\boldsymbol{b} \neq \boldsymbol{0}$. Пусть, кроме того, $\boldsymbol{a} \nparallel \boldsymbol{b}$. Тогда складываем $\boldsymbol{a}$ и $\boldsymbol{b}$ по правилу параллелограмма: откладываем от точки O векторы $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$ и получаем $\overrightarrow{OD} = \boldsymbol{a} + \boldsymbol{b}$ (рис. 53). Если $x > 0$, то векторы $x\boldsymbol{a}$, $x\boldsymbol{b}$ представляются отрезками OA', OB' (где $OA' = x \cdot OA$, $OB' = x \cdot OB$), и получается параллелограмм $OA'D'B'$, подобный первоначальному (рис. 53). Поэтому $OD' = x \cdot OD$, т. е.

$$\overrightarrow{OA}' + \overrightarrow{OB}' = x \cdot \overrightarrow{OD}.$$

Следовательно,

$$x\overrightarrow{OA} + x\overrightarrow{OB} = x(\overrightarrow{OA} + \overrightarrow{OB});$$

таким образом, $x\boldsymbol{a} + x\boldsymbol{b} = x(\boldsymbol{a} + \boldsymbol{b})$ при $x > 0$.

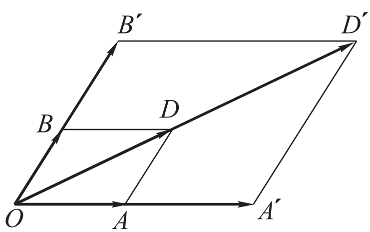


Рис. 53

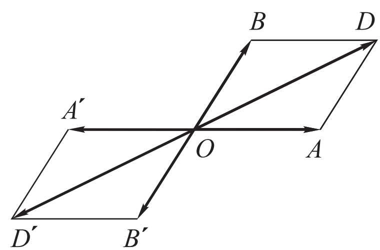


Рис. 54

Если $x < 0$, то нужно только заменить векторы на противоположные и применить гомотетию с коэффициентом $|x|$ (рис. 54). $\square$

§ 4. Скалярное произведение

Угол между векторами. *Углом* между двумя ненулевыми векторами $\boldsymbol{a}$, $\boldsymbol{b}$ называется угол, образованный представляющими их направленными отрезками $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$, отложенными от одной точки.

Лемма. *Этот угол не зависит от точки* O (без этого свойства его нельзя было бы считать углом между векторами).

Действительно, пусть (неколлинеарные) векторы $\boldsymbol{a}$, $\boldsymbol{b}$ отложены от точек O, O', так что (рис. 55)

$$\overrightarrow{OA} = \overrightarrow{O'A'} = \boldsymbol{a}, \quad \overrightarrow{OB} = \overrightarrow{O'B'} = \boldsymbol{b}.$$

Тогда

$$\overrightarrow{AB} = \overrightarrow{A'B'} = \boldsymbol{b} - \boldsymbol{a}.$$

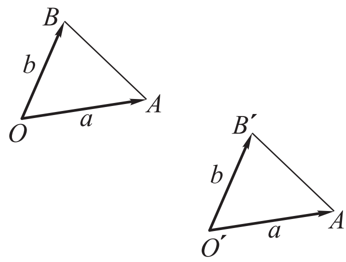


Рис. 55

Поэтому треугольники OAB и $O'A'B'$ равны «по трем сторонам». Стало быть, их соответственные углы равны, так что $\angle O = \angle O'$, что и требовалось доказать. □

Угол между векторами $\boldsymbol{a}$, $\boldsymbol{b}$ обозначается $\angle(\boldsymbol{a},\ \boldsymbol{b})$.

Ортогональность. Два свободных ненулевых вектора называются *ортогональными,* если угол между ними равен 90°, т. е. если представляющие их направленные отрезки, отложенные из одной точки, перпендикулярны. Нулевой вектор считается ортогональным любому. Обозначение — то же, что для перпендикулярности: $\boldsymbol{a} \perp \boldsymbol{b}$ и т. д.

Скалярное произведение. *Скалярным произведением* двух ненулевых векторов называется произведение их длин на косинус угла между ними. Для векторов $\boldsymbol{a}$, $\boldsymbol{b}$ оно обозначается $\boldsymbol{ab}$ (или $\boldsymbol{a} \cdot \boldsymbol{b}$), так что по определению

$$\boldsymbol{ab} = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cdot \cos\angle(\boldsymbol{a},\ \boldsymbol{b}). \tag{1}$$

Если же хотя бы один из векторов $\boldsymbol{a}$, $\boldsymbol{b}$ нулевой, то $\boldsymbol{ab} = 0$.

Отметим два важных частных случая.

1. Если $\boldsymbol{a} = \boldsymbol{b}$, то $\angle(\boldsymbol{a},\ \boldsymbol{b}) = 0$ и $|\boldsymbol{a}| = |\boldsymbol{b}|$. Поэтому $\boldsymbol{a} \cdot \boldsymbol{a} = |\boldsymbol{a}|^2$. Произведение $\boldsymbol{a} \cdot \boldsymbol{a}$ обозначается $\boldsymbol{a}^2$ и называется *скалярным квадратом* или просто *квадратом* вектора.

2. Для ненулевых векторов $\boldsymbol{a}$, $\boldsymbol{b}$ их скалярное произведение равно нулю тогда и только тогда, когда $\cos\angle(\boldsymbol{a}, \boldsymbol{b}) = 0$, т. е. векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ ортогональны: $\boldsymbol{a} \perp \boldsymbol{b}$. Итак, $\boldsymbol{ab} = 0$ равносильно тому, что $\boldsymbol{a} \perp \boldsymbol{b}$.

Алгебраические свойства скалярного умножения (в них $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ — любые векторы и x — любое число).

(I) Коммутативность: $\boldsymbol{ab} = \boldsymbol{ba}$.

(II) Ассоциативность по отношению к числовому множителю: $(x\boldsymbol{a})\boldsymbol{b} = x(\boldsymbol{ab})$.

(III) Дистрибутивность относительно сложения: $\boldsymbol{a}(\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{ab}+\boldsymbol{bc}$.

Доказательство. I. Ясно из определения по формуле (1).

II. Если $x = 0$ при $\boldsymbol{a} = \boldsymbol{0}$ или $\boldsymbol{b} = \boldsymbol{0}$, то равенство верно: обе его чисти — нули.

Пусть $x \neq 0$, $\boldsymbol{a} \neq \mathbf{0}$, $\boldsymbol{b} \neq \mathbf{0}$. Тогда

$$(x\boldsymbol{a})\boldsymbol{b} = |x\boldsymbol{a}||\boldsymbol{b}| \cos\angle(x\boldsymbol{a},\ \boldsymbol{b}) = |x||\boldsymbol{a}||\boldsymbol{b}| \cos\angle(x\boldsymbol{a},\ \boldsymbol{b}). \tag{2}$$

Если $x > 0$, то $|x| = x$ и $x\boldsymbol{a} \uparrow\uparrow \boldsymbol{a}$, так что $\angle(x\boldsymbol{a},\ \boldsymbol{b}) = \angle(\boldsymbol{a},\ \boldsymbol{b})$. Поэтому

$$(x\boldsymbol{a})\boldsymbol{b} = x|\boldsymbol{a}||\boldsymbol{b}| \cos\angle(\boldsymbol{a},\ \boldsymbol{b}) = x(\boldsymbol{a}\boldsymbol{b}).$$

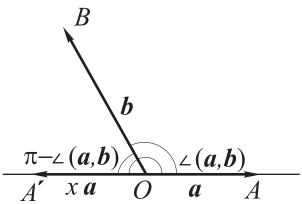


Рис. 56

Если $x < 0$, то $|x| = -x$. Но $x\boldsymbol{a} \uparrow\downarrow \boldsymbol{a}$, так что (рис. 56) $\angle(x\boldsymbol{a},\ \boldsymbol{b}) = \pi - \angle(\boldsymbol{a},\ \boldsymbol{b})$ и $\cos\angle(x\boldsymbol{a},\ \boldsymbol{b}) = -\cos\angle(\boldsymbol{a},\ \boldsymbol{b})$. Поэтому из (2)

$$(x\boldsymbol{a})\boldsymbol{b} = -x|\boldsymbol{a}||\boldsymbol{b}|(-\cos\angle(\boldsymbol{a},\ \boldsymbol{b})) = \\ = x|\boldsymbol{a}||\boldsymbol{b}| \cos\angle(\boldsymbol{a},\ \boldsymbol{b});$$

таким образом, формула (II) верна во всех случаях.

III. Для доказательства дистрибутивности докажем сначала важные формулы:

$$(\boldsymbol{a} + \boldsymbol{b})^2 = \boldsymbol{a}^2 + \boldsymbol{b}^2 + 2\boldsymbol{a}\boldsymbol{b},$$
$$(\boldsymbol{a} - \boldsymbol{b})^2 = \boldsymbol{a}^2 + \boldsymbol{b}^2 - 2\boldsymbol{a}\boldsymbol{b}. \tag{3}$$

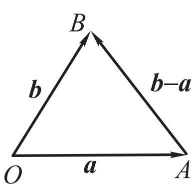


Рис. 57

Действительно, если один из векторов нулевой, то формулы очевидно верны.

Пусть векторы $\boldsymbol{a}$, $\boldsymbol{b}$ не равны нулю. Отложив их от одной точки, получим (рис. 57)

$$\overrightarrow{OA} = \boldsymbol{a}, \quad \overrightarrow{OB} = \boldsymbol{b}, \quad \overrightarrow{AB} = \boldsymbol{b} - \boldsymbol{a}.$$

По известной формуле косинусов

$$AB^2 = OA^2 + OB^2 - 2OA \cdot OB \cdot \cos O$$

или в векторах

$$(\boldsymbol{a} - \boldsymbol{b})^2 = |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 - 2|\boldsymbol{a}| \cdot |\boldsymbol{b}| \cdot \cos\angle(\boldsymbol{a},\ \boldsymbol{b}),$$

и, по определению скалярного произведения,

$$(\boldsymbol{a} - \boldsymbol{b})^2 = \boldsymbol{a}^2 + \boldsymbol{b}^2 - 2\boldsymbol{a}\boldsymbol{b}.$$

Так как $\boldsymbol{a} \cdot (-\boldsymbol{b}) = -\boldsymbol{ab}$ (по свойству II скалярного произведения), то

$$(\boldsymbol{a} + \boldsymbol{b})^2 = \boldsymbol{a}^2 + \boldsymbol{b}^2 + 2\boldsymbol{ab}.$$

Отметим еще, что из доказанных формул (3) непосредственно следует, что

$$(\boldsymbol{a} + \boldsymbol{b})^2 + (\boldsymbol{a} - \boldsymbol{b})^2 = 2(\boldsymbol{a}^2 + \boldsymbol{b}^2). \tag{4}$$

Теперь обратимся непосредственно к доказательству дистрибутивности.

Пусть даны три вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$. Положим

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{p}, \quad \boldsymbol{a} + \boldsymbol{c} = \boldsymbol{q} \tag{5}$$

и применим формулу (4) к векторам $\boldsymbol{p}$, $\mathbf{q}$:

$$(\boldsymbol{p} + \boldsymbol{q})^2 + (\boldsymbol{p} - \boldsymbol{q})^2 = 2(\boldsymbol{p}^2 + \boldsymbol{q}^2).$$

Из (5) находим (пользуясь формулами (3))

$$(\boldsymbol{p} + \boldsymbol{q})^2 = (2\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}))^2 = 4\boldsymbol{a}^2 + (\boldsymbol{b} + \boldsymbol{c})^2 + 4\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}).$$

Кроме того, $(\boldsymbol{p} - \boldsymbol{q})^2 = (\boldsymbol{b} - \boldsymbol{c})^2$, и так как $(\boldsymbol{b} + \boldsymbol{c})^2 + (\boldsymbol{b} - \boldsymbol{c})^2 = 2(\boldsymbol{b}^2 + \boldsymbol{c}^2)$, то

$$(\boldsymbol{p} + \boldsymbol{q})^2 + (\boldsymbol{p} - \boldsymbol{q})^2 = 4\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) + 4\boldsymbol{a}^2 + 2(\boldsymbol{b}^2 + \boldsymbol{c}^2).$$

Вместе с тем, из (5),

$$2(\boldsymbol{p}^2 + \boldsymbol{q}^2) = 2(\boldsymbol{a} + \boldsymbol{b})^2 + 2(\boldsymbol{a} + \boldsymbol{c})^2 = 4\boldsymbol{a}^2 + 2\boldsymbol{b}^2 + 2\boldsymbol{c}^2 + 4\boldsymbol{ab} + 4\boldsymbol{ac}.$$

Сравнивая это с предыдущей формулой, в силу равенства (4) получаем

$$\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{ab} + \boldsymbol{ac},$$

что и требовалось доказать. $\square$

Замечание. Работа силы $\boldsymbol{F}$ на пути $\boldsymbol{s}$ выражается формулой

$$W = |\boldsymbol{F}| \cdot |\boldsymbol{s}| \cdot \cos \angle(\boldsymbol{F},\ \boldsymbol{s}),$$

т. е. $W = \boldsymbol{F} \cdot \boldsymbol{s}$. Пусть на тело действуют две силы $\boldsymbol{F}_1$, $\boldsymbol{F}_2$. Тогда работа их равнодействующей $\boldsymbol{F} = \boldsymbol{F}_1 + \boldsymbol{F}_2$ будет

$$(\boldsymbol{F}_1 + \boldsymbol{F}_2)\boldsymbol{s} = \boldsymbol{F}_1 \cdot \boldsymbol{s} + \boldsymbol{F}_2 \cdot \boldsymbol{s}.$$

То есть работа суммы сил равна сумме их работ: работы складываются вместе с силами.

Диагональ параллелепипеда. Пусть $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$, $\overrightarrow{OC} = \boldsymbol{c}$ — векторы, идущие по ребрам параллелепипеда, и $\overrightarrow{OD} = \boldsymbol{d}$ — вектор диагонали. Тогда $\boldsymbol{d} = \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c}$. Поэтому если a, b, c — длины ребер и α, β, γ — углы между OB и OC и т. д., то

$$\begin{aligned} d^2 = \boldsymbol{d}^2 = (\boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c})^2 = \\ = \boldsymbol{a}^2 + \boldsymbol{b}^2 + \boldsymbol{c}^2 + 2\boldsymbol{ab} + 2\boldsymbol{bc} + 2\boldsymbol{ca} = \\ = a^2 + b^2 + c^2 + 2ab\cos\gamma + 2bc\cos\alpha + 2ca\cos\beta. \end{aligned}$$

Проекция вектора на ось. *Осью* называют прямую, на которой указано направление. Пусть $\boldsymbol{a}$ — данный вектор и e — ось. Представим вектор $\boldsymbol{a}$ отрезком $\overrightarrow{AB} = \boldsymbol{a}$ и спроектируем на ось e. Пусть A', B' — проекции точек A, B. *Проекцией вектора на ось* называют длину отрезка $A'B'$, если направление вектора $\overrightarrow{A'B'}$ совпадает с направлением оси, и длину $A'B'$ с минусом, если направление вектора $\overrightarrow{A'B'}$ противоположно направлению оси (рис. 58).

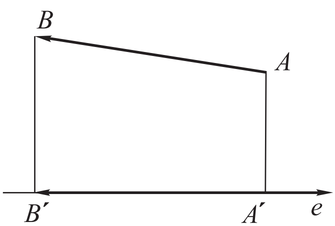


Рис. 58

Направление оси можно указать единичным вектором $\mathbf{e}$. Тогда проекция на ось равна

$$\boldsymbol{ae} = |\boldsymbol{a}|\cos\angle(\boldsymbol{a},\ \boldsymbol{e}).$$

(Докажите это.) Тут заключен, между прочим. тот вывод, что проекция на ось не зависит от выбора точки A, от которой откладывается вектор $\boldsymbol{a}$.

Скалярное произведение в координатах.

Теорема 1. *Если основные векторы единичные и взаимно ортогональные, то скалярное произведение любых двух векторов равно сумме произведений их одноименных координат.*

На плоскости

$$\boldsymbol{ab} = a_1b_1 + a_2b_2. \tag{6}$$

В частности, если $\boldsymbol{a} = \boldsymbol{b}$, то получаем $\boldsymbol{a}^2 = a_1^2 + a_2^2$.

В пространстве

$$\boldsymbol{ab} = a_1b_1 + a_2b_2 + a_3b_3, \tag{7}$$

$$\boldsymbol{a}^2 = a_1^2 + a_2^2 + a_3^2. \tag{8}$$

Доказательство. Пусть основные векторы на плоскости $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ единичные и взаимно ортогональные, т. е.

$$\boldsymbol{e}_1^2 = \boldsymbol{e}_2^2 = 1, \quad \boldsymbol{e}_1\boldsymbol{e}_2 = 0. \tag{9}$$

Векторы $\boldsymbol{a}$, $\boldsymbol{b}$ в той же плоскости представляются через $\boldsymbol{e}_1$, $\boldsymbol{e}_2$:

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2, \quad \boldsymbol{b} = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2.$$

Выполняем умножение, пользуясь свойствами скалярного произведения:

$$\boldsymbol{ab} = (a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2)(b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2) = \\ = a_1b_1\boldsymbol{e}_1^2 + a_1b_2\boldsymbol{e}_1\boldsymbol{e}_2 + a_2b_1\boldsymbol{e}_2\boldsymbol{e}_1 + a_2b_2\boldsymbol{e}_2^2.$$

Отсюда ввиду (9) получаем выражение (6).

Рассмотрим в пространстве три основных вектора $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$. Если они единичные взаимно ортогональные, то

$$\boldsymbol{e}_1^2 = \boldsymbol{e}_2^2 = \boldsymbol{e}_3^2 = 1, \quad \boldsymbol{e}_1\boldsymbol{e}_3 = \boldsymbol{e}_2\boldsymbol{e}_3 = \boldsymbol{e}_3\boldsymbol{e}_1 = 0. \tag{10}$$

Векторы $\boldsymbol{a}$, $\boldsymbol{b}$ в пространстве представляются через эти векторы:

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3, \quad \boldsymbol{b} = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + b_3\boldsymbol{e}_3.$$

Выписав произведение $\boldsymbol{ab}$ и воспользовавшись (10), получим подобно предыдущему выражение (7). □

Если $\boldsymbol{e}$ — единичный вектор, то его проекции на оси или, что то же, скалярные произведения $\boldsymbol{e}_1\boldsymbol{e}$, $\boldsymbol{e}_2\boldsymbol{e}$, $\boldsymbol{e}_3\boldsymbol{e}$ суть не что иное, как косинусы углов α_1, α_2, α_3, которые он образует с осями (поскольку $|\boldsymbol{e}| = 1$ и $|\boldsymbol{e}_1| = |\boldsymbol{e}_2| = |\boldsymbol{e}_3| = 1$):

$$\boldsymbol{e}_1\boldsymbol{e} = \cos\angle(\boldsymbol{e}_1,\ \boldsymbol{e}) = \cos\alpha_1, \\ \boldsymbol{e}_2\boldsymbol{e} = \cos\angle(\boldsymbol{e}_2,\ \boldsymbol{e}) = \cos\alpha_2, \\ \boldsymbol{e}_3\boldsymbol{e} = \cos\angle(\boldsymbol{e}_3,\ \boldsymbol{e}) = \cos\alpha_3.$$

Это так называемые «направляющие косинусы» вектора $\boldsymbol{e}$. Так как $\boldsymbol{e}^2 = 1$, то

$$\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3 = 1.$$

§ 5. Координаты в пространстве

Прямоугольные координаты. Забыв пока о векторах, опишем, как вводятся в пространстве прямоугольные координаты.

Возьмем какую-нибудь плоскость α и введем на ней прямоугольные координаты x, y. Любой точке M в пространстве отнесем три координаты: две координаты x, y — это координаты ее проекции M_α на плоскость α, а третью координату z определяем так: $|z|$ равно расстоянию точки M от плоскости α, причем $z > 0$ с одной какой-нибудь стороны от нее, $z < 0$ с другой стороны и $z = 0$ на самой α (рис. 59). Так как расстояние от точки до плоскости равно длине перпендикуляра, опущенного из точки на плоскость, то $|z| = MM_\alpha$.

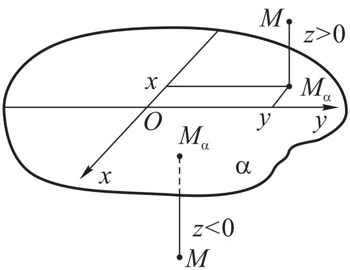


Рис. 59

Так каждой точке пространства однозначно относятся три координаты x, y, z: координата x считается первой, y — второй, z — третьей.

Обратно: *если заданы любые три числа в определенном порядке* x_0, y_0, z_0, *то найдется точка* M *с координатами* $x = x_0$, $y = y_0$, $z = z_0$.

Действительно, берем на плоскости α точку M_α с координатами $x = x_0$, $y = y_0$. Если $z_0 = 0$, то это и есть точка M. Иначе проводим из M_α перпендикуляр к плоскости α на длину $|z_0|$ в ту сторону, где $z > 0$, если $z_0 > 0$, и в обратную сторону, если $z_0 < 0$. Конец M этого перпендикуляра и будет иметь координаты x_0, y_0, z_0. □

В изложенном определении прямоугольных координат координата z занимает особое положение. Однако нетрудно определить координаты так, чтобы все три играли одинаковую роль.

Выберем в пространстве какую-нибудь точку O и проведем через нее три взаимно перпендикулярные прямые. Введем на них координаты с общим началом в точке O (рис. 60) и назовем эти координаты x, y, z. Проведенные прямые называются *осями координат:* «ось x», «ось y», «ось z». «Плоскостью xy» называется плоскость, проходящая через оси x и y. Аналогично определяются плоскости xz и yz. (В предыдущем определении координат плоскостью xy была исходная плоскость α.)

Определим координаты любой точки M следующим образом. Берем ее проекции M_x, M_y, M_z на оси и координаты их на осях прини-

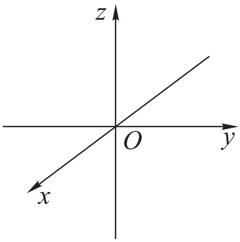


Рис. 60

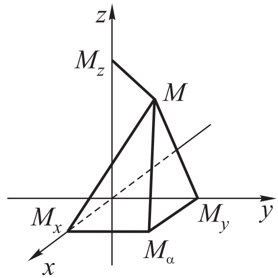


Рис. 61

маем за координаты x, y, z точки M (рис. 61). (Убедитесь, что это те же самые координаты, какие были определены сначала.)

Принято изображать координатные оси так, как на рис. 61. Ось z считается вертикальной, направленной вверх; оси x, y считаются горизонтальными: ось x — направленной вперед (на зрителя), ось y — вправо. Такая система координат называется *правой*. Если представить себе винт, ввинчивающийся в положительном направлении по оси z, то головка винта должна вращаться от положительной полуоси x к положительной полуоси y. Если изменить направление оси x (или любой другой) на противоположное, то получится «левая» система координат. Ей соответствует не обычный, а левый винт.

Нахождение координат точки и, обратно, нахождение точки с данными координатами проще всего представлять так. Опускаем из M перпендикуляр MM_α на плоскость xy, из точки M_α опускаем перпендикуляр $M_\alpha M_x$ на ось x. Длины отрезков, взятые с должными знаками, и дадут координаты точки M: MM_α дает z, $M_\alpha M_x$ дает y, M_xO дает x.

Точку с данными координатами x_0, y_0, z_0 находим построением в обратном порядке: сначала строим точку M_x, затем по перпендикуляру к оси x — точку M_α и, наконец, по перпендикуляру к плоскости xy находим $M(x_0,\ y_0,\ z_0)$.

Координаты и векторы. Радиус-вектор точки. Систему координат можно задать, указав ее начало O и единичные векторы по осям. Это одинаково возможно на плоскости и в пространстве; различие лишь в том, что на плоскости задаются два вектора, а в пространстве — три. Если строится прямоугольная система координат, то векторы берутся взаимно ортогональными. Рассмотрим введение ко-

ординат в пространстве (на плоскости все будет аналогично и только несколько проще).

Выберем в пространстве точку O и три взаимно ортогональных единичных вектора $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$:

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = 1, \quad \boldsymbol{ij} = \boldsymbol{jk} = \boldsymbol{ki} = 0. \tag{1}$$

Каждая точка M оказывается концом вектора $\overrightarrow{OM}$. Он называется *радиус-вектором* точки M. С другой стороны, всякий вектор $\boldsymbol{r}$, если его отложить от точки O, задает некоторую точку M — конец радиус-вектора $\overrightarrow{OM} = \boldsymbol{r}$. Вектор $\boldsymbol{r}$ определяется своими координатами x, y, z относительно основных векторов $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$. Вместе с ним определяется и точка M:

$$\overrightarrow{OM} = \boldsymbol{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z. \tag{2}$$

Координаты радиус-вектора $\overrightarrow{OM}$ и принимаются за координаты точки M.

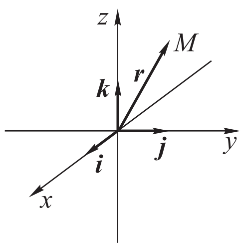


Рис. 62

Убедимся, что это те же координаты, какие были определены в предыдущем пункте. Отложим векторы $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ от начала O, получим «трехвекторник» $O\boldsymbol{ijk}$ из векторов $\overrightarrow{OA} = \boldsymbol{i}$, $\overrightarrow{OB} = \boldsymbol{j}$, $\overrightarrow{OC} = \boldsymbol{k}$. Проведем прямые OA, OB, OC — это будут оси координат x, y, z; направление на них указывается векторами $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ (рис. 62).

Из равенства (2), умножая на $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, находим

$$x = \boldsymbol{ir}, \quad y = \boldsymbol{jr}, \quad z = \boldsymbol{kr}, \tag{3}$$

т. е. x, y, z — это проекции радиус-вектора $\boldsymbol{r}$ на оси координат. Если M_x, M_y, M_z — проекции точки M на оси, то проекции на оси вектора $\boldsymbol{r} = \overrightarrow{OM}$ — это длины отрезков OM_x, OM_y, OM_z с соответствующим знаком. Если, например, M_x на положительной полуоси, то направление $\overrightarrow{OM}_x$ совпадает с направлением оси x и проекция на ось x, по определению, положительна; в противном случае она отрицательна. Поэтому координаты проекций точки M на осях — это те же самые x, y, z, что находятся по формулам (3).

Общие прямолинейные (аффинные) координаты. Можно взять любые три некомпланарных вектора $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$. Вместе с точкой O они образуют трехвекторник $O\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3$ (рис. 63). Радиус-вектор

представляется в виде

$$\overrightarrow{OM} = \boldsymbol{r} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3.$$

Числа x_1, x_2, x_3 задают радиус-вектор и тем самым — его конец, точку M, т. е. они служат ее координатами относительно трехвекторника $O\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3$.

Таким образом в пространстве вводятся координаты относительно любого трехвекторника; аналогично на плоскости вводятся координаты относительно любого двухвекторника.

Геометрически эти координаты определяются так. Векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ откладываются от начала O; получаем $\overrightarrow{OA} = \boldsymbol{e}_1$, $\overrightarrow{OB} = \boldsymbol{e}_2$, $\overrightarrow{OC} = \boldsymbol{e}_3$. Проводим прямые OA, OB, OC — это будут оси x_1, x_2, x_3; на них задаем направления по векторам $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ и тем самым определяем положительные и отрицательные полуоси (рис. 63). На осях вводим координаты: если M — точка на оси x_1, то ее координата x_1 определяется из формулы

$$\overrightarrow{OM} = x_1\boldsymbol{e}_1.$$

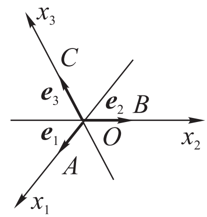
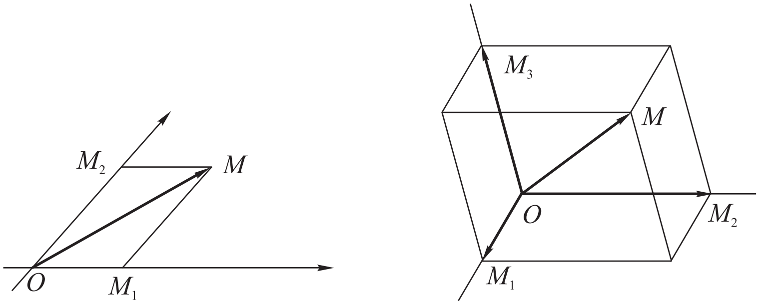


Рис. 63

То есть если $\overrightarrow{OM} \uparrow\uparrow \boldsymbol{e}_1$, то $x_1 = |OM|/|\boldsymbol{e}_1|$; если же $\overrightarrow{OM} \uparrow\downarrow \boldsymbol{e}_1$, то $x_1 = -|OM|/|\boldsymbol{e}_1|$. Другими словами, если точка M лежит на положительной полуоси x_1, то $x_1 > 0$; если же на отрицательной полуоси, то $x_1 < 0$. А $|x_1|$ — это, можно сказать, длина отрезка OM, измеренная в масштабе $|\boldsymbol{e}_1|$. Так же определяются координаты на осях x_2, x_3 (т. е. аналогично определению координат при единичных векторах на осях).

Координаты точки M в пространстве находим, проектируя ее на оси параллельно координатным плоскостям. Именно, проекция на ось M_1 — это такая точка на оси x_1, что прямая MM_1 параллельна плоскости x_2x_3. Это равносильно тому, что M_1 — это точка, в которой плоскость, проходящая через M параллельно плоскости x_2x_3, пересекает ось x_1. Аналогично находятся проекции M_2, M_3 точки M на оси x_2, x_3.

Отрезки OM_1, OM_2, OM_3 на осях служат ребрами параллелепипеда с диагональю OM (рис. 64, *б*). (Особые случаи, когда точка M лежит на одной из координатных плоскостей или на какой-либо оси, легко рассматриваются дополнительно.)



а *б*

Рис. 64

Аналогично, на плоскости координаты x_1, x_2 определяются для точки M из разложения ее радиус-вектора:

$$\overrightarrow{OM} = \boldsymbol{r} = \boldsymbol{e}_1 x_1 + \boldsymbol{e}_2 x_2.$$

Оси координат проводятся через O вдоль векторов $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ и координаты на них определяются как и выше. Координаты точки M определяются по ее проекциям M_1, M_2 вдоль прямых, параллельных осям. Вектор $\overrightarrow{OM}$ служит диагональю параллелограмма со сторонами $\overrightarrow{OM}$, $\overrightarrow{OM}_2$ (рис. 64, *а*).

Читателю стоит рассмотреть эти построения на плоскости и в пространстве: как определяют координаты точки и как по данным координатам находят точку.

Введенные здесь координаты на плоскости и в пространстве называют *аффинными* (в случае плоскости они упоминались в гл. I, § 4). Их называют *косоугольными* (в отличие от прямоугольных), когда векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ единичные, т. е. когда на осях обычный масштаб, но они не взаимно перпендикулярны.

Вектор в системе координат. При данных основных векторах всякий вектор задается своими координатами. Пусть вектор $\boldsymbol{a}$ отложен от точки A, так что получаем направленный отрезок $\overrightarrow{AB} = \boldsymbol{a}$. Точки A, B задаются радиус-векторами $\overrightarrow{OA}$, $\overrightarrow{OB}$:

$$\overrightarrow{OA} = \boldsymbol{e}_1 a_1 + \boldsymbol{e}_2 a_2 + \boldsymbol{e}_3 a_3,$$
$$\overrightarrow{OB} = \boldsymbol{e}_1 b_1 + \boldsymbol{e}_2 b_2 + \boldsymbol{e}_3 b_3,$$

где $a_1, \ \ldots, \ b_3,$ — координаты точек A, B. Так как $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$, то

$$\overrightarrow{AB} = \boldsymbol{e}_1(b_1 - a_1) + \boldsymbol{e}_2(b_2 - a_2) + \boldsymbol{e}_3(b_3 - a_3).$$

То есть мы получаем результат:

Координаты вектора, представленного направленным отрезком, равны разностям координат его конца и начала. □

§ 6. Правые и левые тройки векторов. Векторное произведение. Смешанное произведение

Упорядоченной тройкой или просто *тройкой* векторов называется совокупность трех векторов, в которой указан их порядок: какой является первым, какой — вторым, какой — третьим.

При записи тройки векторов они всегда располагаются в порядке их номеров.

Тройка некомпланарных векторов называется *правой (левой)*, если эти векторы, отложенные от одного начала, располагаются так же, как расставленные (примерно под прямыми углами) пальцы правой (левой) руки: большой палец — по первому вектору, указательный — по второму, средний — по третьему (рис. 65).

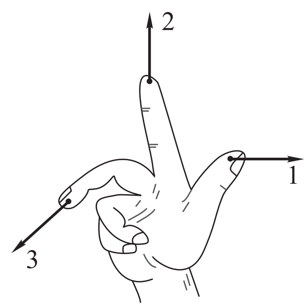


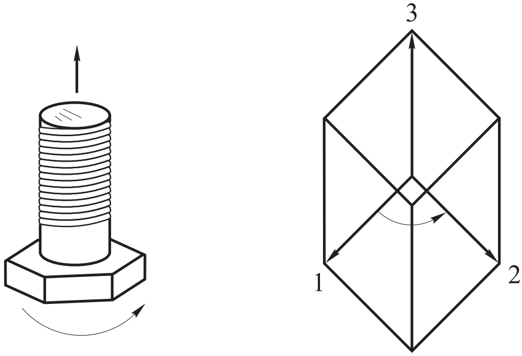
Рис. 65

Тройку векторов, отложенных от одной точки — общего их начала, — будем называть *трехвекторником*. Если векторы компланарны, трехвекторник назовем *плоским*. Соответственно только что данному определению различаются правые и левые трехвекторники. Так как векторы всегда можно отложить от одной точки, то говорим ли мы здесь о тройках или трехвекторниках — безразлично. Точно так же система прямолинейных координат называется *правой* (*левой*), если основные ее векторы образуют правую (левую) тройку.

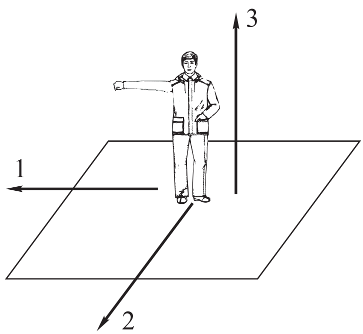
Указанное правило различения правых и левых троек равносильно каждому из следующих четырех.

1. Трехвекторник правый, если третий вектор направлен от плоскости двух первых в ту сторону (в то полупространство), куда движется правый винт, когда его головка вращается от первого вектора ко второму (рис. 66, *а*) Если то же будет для левого винта, то трехвекторник левый.

2. Трехвекторник правый, если при взгляде на него изнутри построенного на нем параллелепипеда (из первого квадранта построенной на нем системы координат) векторы видны так, что переход от первого ко второму и дальше — к третьему — происходит против часовой стрелки (рис. 66, *б*). Если же это переход по часовой стрелке, то трехвекторник левый.



а *б*



в

Рис. 66

3. Трехвекторник правый, если при взгляде на плоскость первых двух векторов со стороны третьего поворот от первого ко второму виден как идущий против часовой стрелки. Если — по часовой стрелке, то трехвекторник левый.

4. Трехвекторник правый, если после того, как вы встали на плос-

кость двух первых векторов со стороны третьего и протянули правую руку по первому вектору, второй вектор оказался направленным вперед (рис. 66, *в*). У левого трехвекторника — то же для левой руки.

(Придумайте еще правила для правого (левого) трехвекторника.)

Различение правых и левых троек векторов имеет важное значение в физике — в законах электромагнетизма. Например, направление напряжения магнитного поля вокруг проводника с током определяется «по правилу правой руки».

Но в геометрии самой по себе нет ни правого, ни левого, — ни пальцев, ни винтов, ни часовых стрелок. Поэтому данные наглядные определения правых и левых троек не относятся к самой геометрии. В ней должно быть принято определение, основанное только на ее собственных понятиях; как его можно дать, мы сейчас скажем.

Определение. Выберем какой-нибудь прямоугольный трехвекторник и назовем его *основным.* *Правыми* будем называть трехвекторники, которые можно получить из основного непрерывным движением и изменением векторов без того, чтобы трехвекторник становился плоским. В противном случае трехвекторник назовем левым.

При этом можно доказать, что левые трехвекторники также переводимы друг в друга, непрерывно и без обращения в плоские. Если две тройки векторов либо обе правые, либо обе левые, то говорят, что они *одной ориентации.* Если же одна правая, другая левая, то говорят, что они *разной ориентации.*

Если основной трехвекторник представить пальцами правой руки, то мы получим наглядное различие правых и левых трехвекторников. Но в самой геометрии понятие правое и левое число условны, реально в ней только то, что есть два вида трехвекторников — две их разные ориентации.

Векторное произведение.

Определение. *Векторным произведением* вектора $\boldsymbol{a}$ на не коллинеарный ему вектор $\boldsymbol{b}$ называют вектор, обозначаемый $\boldsymbol{a} \times \boldsymbol{b}$, который определяется следующими тремя условиями:

1) длина вектора $\boldsymbol{a} \times \boldsymbol{b}$ равна произведению длин векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ на синус угла между ними:

$$|\boldsymbol{a} \times \boldsymbol{b}| = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cdot \sin \angle(\boldsymbol{a},\ \boldsymbol{b}).$$

2) вектор $\boldsymbol{a} \times \boldsymbol{b}$ ортогонален обоим векторам $\boldsymbol{a}$, $\boldsymbol{b}$;

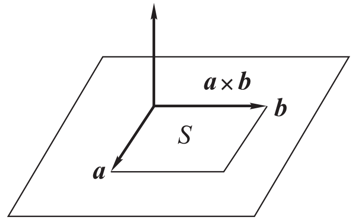


Рис. 67

3) вектор $\boldsymbol{a} \times \boldsymbol{b}$ направлен так, что векторы $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{a} \times \boldsymbol{b}$ образуют правую тройку (рис. 67).

Если векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ коллинеарны, то положим $\boldsymbol{a} \times \boldsymbol{b} = 0$.

Модуль векторного произведения $\boldsymbol{a} \times \boldsymbol{b}$ равен площади параллелограмма, сторонами которого служат векторы $\boldsymbol{a}$, $\boldsymbol{b}$, отложенные от одной точки. Это следует из того, что площадь S такого параллелограмма как раз равна $|\boldsymbol{a}||\boldsymbol{b}|\sin\angle(\boldsymbol{a},\ \boldsymbol{b})$.

Если векторы $\boldsymbol{a}$, $\boldsymbol{b}$ неколлинеарны, то, отложенные от одной точки, они определяют содержащую их плоскость, и вектор $\boldsymbol{a}\times\boldsymbol{b}$, ортогональный этим векторам, перпендикулярен этой плоскости.

Ввиду этих двух замечаний можно векторное произведение геометрически определить так.

Векторное произведение $\boldsymbol{a} \times \boldsymbol{b}$ *неколлинеарных векторов* $\boldsymbol{a}$, $\boldsymbol{b}$ *по модулю равно площади построенного на них параллелограмма, перпендикулярно его плоскости, и образует с* $\boldsymbol{a}$, $\boldsymbol{b}$ *правую тройку.*

Из определения векторного произведения следует:

Векторное произведение равно нулю тогда и только тогда, когда сомножители коллинеарны.

Алгебраические свойства векторного произведения (в них $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ — любые векторы и λ — любое число).

1. Антикоммутативность: $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$.

2. Ассоциативность по отношению к численному множителю:

$$\lambda(\boldsymbol{a} \times \boldsymbol{b}) = (\lambda\boldsymbol{a}) \times \boldsymbol{b}.$$

3. Дистрибутивность относительно сложения:

$$\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}.$$

Первые два свойства легко устанавливаются из определения. Перестановка векторов $\boldsymbol{a}$, $\boldsymbol{b}$ изменяет направление вращения от первого сомножителя ко второму. О втором свойстве можно заметить, что оно очевидно при $\lambda > 0$, а при $\lambda < 0$ вектор $\lambda\boldsymbol{a}$ направлен противоположно $\boldsymbol{a}$ и вращение от $\lambda\boldsymbol{a}$ к $\boldsymbol{b}$ противоположно вращению от $\boldsymbol{a}$ к $\boldsymbol{b}$.

Трудность представляет доказательство третьего свойства — дистрибутивности. Для его доказательства заметим следующее: если вектор $\boldsymbol{a}$ единичный и $\boldsymbol{b}$ ему ортогонален, то вектор $\boldsymbol{a} \times \boldsymbol{b}$ получится, если повернуть $\boldsymbol{b}$ в плоскости, перпендикулярной $\boldsymbol{a}$, на прямой угол в таком направлении, чтобы он образовывал с $\boldsymbol{a}$ и $\boldsymbol{b}$ правую тройку (наглядно: при взгляде на плоскость со стороны $\boldsymbol{a}$ поворот должен быть виден как происходящий против часовой стрелки).

Лемма 1. *Пусть $\boldsymbol{a}$, $\boldsymbol{b}$ — два неколлинеарных вектора, вектор $\boldsymbol{b}_1$ получен проектированием вектора $\boldsymbol{b}$ на плоскость α, перпендикулярную $\boldsymbol{a}$. Утверждается, что $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{b}_1$. При этом векторное произведение $\boldsymbol{a} \times \boldsymbol{b}_1$ направлено в плоскости α по лучу, перпендикулярному $\boldsymbol{b}_1$, так что тройка $\boldsymbol{b}_1$, $\boldsymbol{a} \times \boldsymbol{b}_1$, $\boldsymbol{a}$ правая (наглядно: при повороте винта от $\boldsymbol{b}_1$ к $\boldsymbol{a} \times \boldsymbol{b}_1$ он движется по $\boldsymbol{a}$).*

Доказательство. Площади параллелограммов, натянутых на векторы $\boldsymbol{a}$, $\boldsymbol{b}$ и $\boldsymbol{a}$, $\boldsymbol{b}_1$, равны, так как $\boldsymbol{b}_1$ — высота параллелограмма (рис. 68). Направление же вектора $\boldsymbol{a} \times \boldsymbol{b}_1$, очевидно, то же, что у $\boldsymbol{a} \times \boldsymbol{b}$. Оно перпендикулярно плоскости векторов $\boldsymbol{a}$, $\boldsymbol{b}$; вращение же от $\boldsymbol{a}$ к $\boldsymbol{b}$ и от $\boldsymbol{a}$ к $\boldsymbol{b}_1$ идет в одном направлении. □

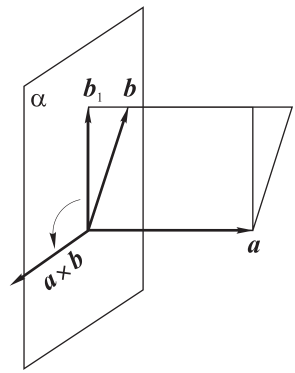


Рис. 68

Лемма 2. *Векторное произведение $\boldsymbol{a} \times \boldsymbol{b}$ можно получить следующим образом:*

а) *проектируем $\boldsymbol{b}$ на плоскость α, перпендикулярную $\boldsymbol{a}$: пусть $\boldsymbol{b}_1$ — эта проекция;*

б) *умножаем полученную проекцию — вектор $\boldsymbol{b}_1$ — на $|\boldsymbol{a}|$;*

в) *полученный вектор $|\boldsymbol{a}| \cdot \boldsymbol{b}_1$ поворачиваем на 90° в плоскости α в направлении, при котором винт идет вдоль $\boldsymbol{a}$.*

Полученный вектор и будет векторным произведением $\boldsymbol{a} \times \boldsymbol{b}$.

Действительно, по лемме 1, $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{b}_1$ и направление вектора $\boldsymbol{a} \times \boldsymbol{b}_1$ определяется как сказано здесь, в лемме 2. Его длина $|\boldsymbol{a} \times \boldsymbol{b}_1|$ равна $|\boldsymbol{a}||\boldsymbol{b}_1|$, поскольку $\boldsymbol{b}_1 \perp \boldsymbol{a}$. □

Теперь докажем дистрибутивность:

$$\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}.$$

Будем получать стоящие здесь произведения, как указано в лемме 2. При этом:

а) проекция суммы $\boldsymbol{b} + \boldsymbol{c}$ равна сумме проекций:

$$(\boldsymbol{b} + \boldsymbol{c})_1 = \boldsymbol{b}_1 + \boldsymbol{c}_1;$$

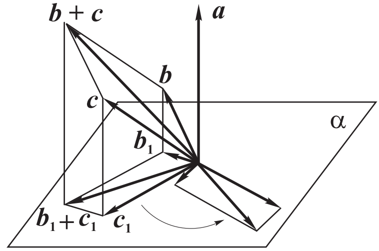


Рис. 69

б) умножение на $|\boldsymbol{a}|$ дает

$$|\boldsymbol{a}|(\boldsymbol{b}+\boldsymbol{c})_1 = |\boldsymbol{a}|\boldsymbol{b}_1 + |\boldsymbol{a}|\boldsymbol{c}_1; \tag{1}$$

в) при повороте взаимное расположение векторов сохраняется, так что сумма переходит в сумму (диагональ параллелограмма поворачивается вместе со сторонами, (см. рис. 69).

Следовательно, из (1) получаем

$$\boldsymbol{a}\times(\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{a}\times\boldsymbol{b} + \boldsymbol{a}\times\boldsymbol{c}. \quad \square$$

Выражение векторного произведения в координатах. Пусть, как принято, $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ обозначают основные векторы правой системы координат, так что они 1) единичные, 2) взаимно перпендикулярные, 3) образуют правую тройку. Тогда из определения векторного произведения непосредственно следует, что

$$\begin{array}{lll} \boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k}, & \boldsymbol{j}\times\boldsymbol{k}=\boldsymbol{i}, & \boldsymbol{k}\times\boldsymbol{i}=\boldsymbol{j}, \\ \boldsymbol{j}\times\boldsymbol{i}=-\boldsymbol{k}, & \boldsymbol{k}\times\boldsymbol{j}=-\boldsymbol{i}, & \boldsymbol{i}\times\boldsymbol{k}=-\boldsymbol{j}, \\ \boldsymbol{i}\times\boldsymbol{i}=\boldsymbol{0}, & \boldsymbol{j}\times\boldsymbol{j}=\boldsymbol{0}, & \boldsymbol{k}\times\boldsymbol{k}=\boldsymbol{0}. \end{array}$$

Пользуясь этими равенствами и правилами действия с векторным произведением, найдем выражение для произведения любых двух векторов:

$$\boldsymbol{a}\times\boldsymbol{b} = (\boldsymbol{i}a_1+\boldsymbol{j}a_2+\boldsymbol{k}a_3)\times(\boldsymbol{i}b_1+\boldsymbol{j}b_2+\boldsymbol{k}b_3).$$

Согласно правилам действия перемножаем эти суммы почленно, но помним о порядке сомножителей (их нельзя переставлять). Так как $\boldsymbol{i}\times\boldsymbol{i}=\boldsymbol{j}\times\boldsymbol{j}=\boldsymbol{k}\times\boldsymbol{k}=\boldsymbol{0}$, то получаем

$$\boldsymbol{a}\times\boldsymbol{b} = (\boldsymbol{i}\times\boldsymbol{j})a_1b_2 + (\boldsymbol{j}\times\boldsymbol{i})a_2b_1 + (\boldsymbol{i}\times\boldsymbol{k})a_1b_3 + \\ + (\boldsymbol{k}\times\boldsymbol{i})a_3b_1 + (\boldsymbol{j}\times\boldsymbol{k})a_2b_3 + (\boldsymbol{k}\times\boldsymbol{j})a_3b_2.$$

Располагаем члены по порядку векторов $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ и тогда получаем

$$\boldsymbol{a}\times\boldsymbol{b} = \boldsymbol{i}(a_2b_3-a_3b_2) + \boldsymbol{j}(a_3b_1-a_1b_3) + \boldsymbol{k}(a_1b_2-a_2b_1).$$

Смешанное (скалярно-векторное) произведение. *Смешанным*, или *скалярно-векторным, произведением* тройки векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ называется скалярное произведение первого вектора на векторное произведение второго на третий:

$$(\boldsymbol{abc}) = \boldsymbol{a}(\boldsymbol{b}\times\boldsymbol{c}).$$

(Порядок сомножителей важен, так как $\boldsymbol{c} \times \boldsymbol{b} = -(\boldsymbol{b} \times \boldsymbol{c})$; но вместе с тем $\boldsymbol{a}(\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{b} \times \boldsymbol{c})\boldsymbol{a}$.)

Теорема 1. *Смешанное произведение* $(\boldsymbol{abc}) = \boldsymbol{0}$ *тогда и только тогда, когда векторы* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ *компланарны. Если же они не компланарны, то* $(\boldsymbol{abc})$ *равно по модулю объему параллелепипеда с ребрами* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ (*отложенными из одной точки*); $(\boldsymbol{abc}) > 0$, *если тройка* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ *правая,* $(\boldsymbol{abc}) < 0$, *если она левая.*

Доказательство. Докажем сначала вторую часть теоремы как более содержательную.

Пусть $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ — правая тройка векторов; мы их откладываем от одной точки. По определению

$$(\boldsymbol{abc}) = \boldsymbol{a}(\boldsymbol{b} \times \boldsymbol{c}) = |\boldsymbol{a}||\boldsymbol{b} \times \boldsymbol{c}| \cos\varphi, \tag{2}$$

где φ — угол между векторами $\boldsymbol{a}$ и $\boldsymbol{b} \times \boldsymbol{c}$. Построим на векторах $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ параллелепипед. Одна из его граней — параллелограмм со сторонами $\boldsymbol{b}$, $\boldsymbol{c}$. Примем ее за основание. По свойству векторного произведения ее площадь равна (рис. 70)

$$S = |\boldsymbol{b} \times \boldsymbol{c}|. \tag{3}$$

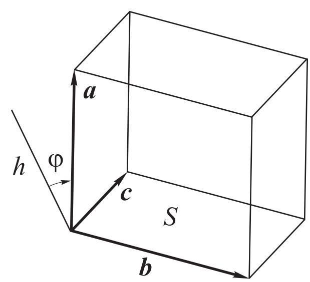


Рис. 70

Другое ребро $\boldsymbol{a}$ образует с перпендикуляром к плоскости основания, направленным как $\boldsymbol{b} \times \boldsymbol{c}$, угол φ. И так как тройка $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ правая, то вектор $\boldsymbol{a}$ направлен в ту же сторону, что $\boldsymbol{b} \times \boldsymbol{c}$. Поэтому угол φ острый, $\cos\varphi > 0$ и

$$|\boldsymbol{a}| \cos\varphi = h \tag{4}$$

— это высота параллелепипеда.

Объем параллелепипеда вычисляется по формуле

$$V = Sh,$$

и из формул (3), (4) получаем

$$V = |\boldsymbol{b} \times \boldsymbol{c}||\boldsymbol{a}| \cos\varphi.$$

Сравнивая с (2), получаем

$$(\boldsymbol{abc}) = V,$$

что и требовалось доказать для правой тройки.

Если тройка векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ левая, то векторы $\boldsymbol{a}$ и $\boldsymbol{b} \times \boldsymbol{c}$ направлены в разные стороны от плоскости векторов $\boldsymbol{b}$, $\boldsymbol{c}$. Поэтому угол φ тупой, $\cos\varphi < 0$, и вместо (4) будет $|\boldsymbol{a}|\cos\varphi = -h$. Поэтому получается, что $(\boldsymbol{abc}) = -V$.

Итак, вторая часть теоремы доказана. Докажем первую ее часть.

Допустим, что $(\boldsymbol{abc}) = 0$, но векторы $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ некомпланарны. Тогда по доказанному было бы $(\boldsymbol{abc}) = \pm V$, где V — объем построенного на них параллелепипеда, т. е. было бы $(\boldsymbol{abc}) \neq 0$. Значит, если $(\boldsymbol{abc}) = 0$, то векторы компланарны.

Пусть теперь векторы $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ компланарны. Если $\boldsymbol{b}$, $\boldsymbol{c}$ коллинеарны, то $\boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{0}$ и $(\boldsymbol{abc}) = 0$; если $\boldsymbol{a} = \boldsymbol{0}$, то опять $(\boldsymbol{abc}) = 0$. Если же $\boldsymbol{a} \neq \boldsymbol{0}$ и $\boldsymbol{b}$, $\boldsymbol{c}$ неколлинеарны, так что $\boldsymbol{b} \times \boldsymbol{c} \neq \boldsymbol{0}$, но все три вектора компланарны, то, значит, $\boldsymbol{a} \perp \boldsymbol{b} \times \boldsymbol{c}$ и $\boldsymbol{a}(\boldsymbol{b} \times \boldsymbol{c}) = 0$, что и требовалось доказать.

Таким образом, теорема доказана полностью. $\square$

В определении произведения $(\boldsymbol{abc})$ векторы играют разную роль, но у параллелепипеда, на них построенного, они — ребра и, стало быть, равноправны. Поэтому из доказанной теоремы вытекает

Следствие 1.

$$(\boldsymbol{abc}) = (\boldsymbol{bca}) = (\boldsymbol{cab}). \tag{5}$$

Доказательство. В доказательстве теоремы 1 применительно к произведению $(\boldsymbol{bca})$ параллелепипед будет тот же, и только за основание будет принята другая грань, построенная на векторах $\boldsymbol{c}$, $\boldsymbol{a}$. Поэтому $(\boldsymbol{bca}) = \pm V$, и именно $(\boldsymbol{bca}) = V$, так как тройка $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{a}$ правая, если $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ правая. То же будет для тройки $\boldsymbol{c}$, $\boldsymbol{b}$, $\boldsymbol{a}$. $\square$

Так как $\boldsymbol{c} \times \boldsymbol{b} = -(\boldsymbol{b} \times \boldsymbol{c})$, то

$$(\boldsymbol{acb}) = \boldsymbol{a}(\boldsymbol{c} \times \boldsymbol{b}) = -\boldsymbol{a}(\boldsymbol{b} \times \boldsymbol{c}) = -(\boldsymbol{abc})$$

и тройка $\boldsymbol{a}$, $\boldsymbol{c}$, $\boldsymbol{b}$ левая, если $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ правая. Отсюда и из равенств (5) вытекает

Следствие 2. *При перестановке любых двух векторов-сомножителей произведение* $(\boldsymbol{abc})$ *меняет знак.* $\square$

Выражение смешанного произведения в координатах. Пусть в некоторой правой прямоугольной системе координаты векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ будут (a_1, a_2, a_3) и т. д. Тогда, пользуясь выражением в координатах для скалярного произведения и векторного произведения, по формуле (2) получим

$$(\boldsymbol{abc}) = \boldsymbol{a}(\boldsymbol{b} \times \boldsymbol{c}) =$$

$$= a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1). \quad (6)$$

Стоящая справа величина есть не что иное, как определитель:

$$(\boldsymbol{abc}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Действительно, разлагая этот определитель по элементам первой строки, мы получим формулу (6).

Таким образом, *смешанное произведение векторов равно определителю из их координат* (относительно основной правой прямоугольной системы координат). □

Об ориентации. Если тройка $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ правая, то $(\boldsymbol{abc}) > 0$, если левая, то $(\boldsymbol{abc}) < 0$, и вследствие (6) это равносильно тому, что такие же знаки имеет определитель из координат вектора.

Это можно превратить в определение, которое будет выглядеть так:

Пусть в пространстве выбрана система прямоугольных координат, которая принята за основную. Тогда тройка некомпланарных векторов называется *правой*, если определитель из их координат относительно основной системы положителен; в противном случае тройка называется *левой*. (Подразумевается, что в определителе порядок строк соответствует порядку векторов в тройке, и порядок столбцов — порядку координат в основной системе.) Основная система координат — тройка ее основных векторов — правая, так как определитель из их координат относительно них самих равен единице:

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Данное формальное определение «правого» и «левого» дает ясное алгебраическое основание для их различения. При этом очевидно, что непрерывным изменением векторов нельзя правую тройку перевести в левую, не делая «по пути» векторы компланарными, так как нельзя от положительных значений определителя перейти к отрицательным, не проходя через нуль, если элементы определителя изменяются непрерывно.

Несколько сложнее доказать, что всякую правую (как и левую) тройку можно непрерывно перевести в любую другую правую (соответственно — в левую), не делая векторы компланарными.

Глава IV

СФЕРА, ПРЯМАЯ, ПЛОСКОСТЬ

§ 1. Расстояние между точками. Сфера. Плоскость

Теорема 1. *В прямоугольных координатах* x, y, z *расстояние между точками* $A(x_A,\ y_A,\ z_A)$ и $B(x_B,\ y_B,\ z_B)$ *выражается как корень квадратный из суммы квадратов разностей их координат:*

$$|AB| = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}. \tag{1}$$

Доказательство. Квадрат вектора $\overrightarrow{AB}$ равен сумме квадратов его координат (в прямоугольной системе координат, когда основные векторы единичные и взаимно перпендикулярные). Вместе с тем координаты вектора $\overrightarrow{AB}$ равны разностям координат его конца и начала: $x_B - x_A$ и т. д. Из этих двух фактов непосредственно следует формула (1). □

Уравнение сферы. Сфера есть геометрическое место (множество) точек, удаленных на заданное расстояние от одной точки — центра сферы; их расстояние от него — это радиус сферы. Таким образом, если O — центр, а R — радиус сферы, то она является множеством всех таких точек M, что $OM^2 = R^2$.

В прямоугольных координатах расстояние выражается формулой (1). Поэтому если x_0, y_0, z_0 — координаты центра O и x, y, z — координаты произвольной (переменной) точки M сферы, то

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2. \tag{2}$$

Это уравнение сферы: ему удовлетворяют координаты всех ее точек и никаких других.

Если возвести уравнение (2) в квадрат и раскрыть скобки, то оно примет вид

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0. \tag{3}$$

Аналогично вопросу, рассмотренному в связи с уравнением окружности, можно поставить вопрос: при каких условиях на коэффициенты $a,\ \ldots,\ d$ уравнение (3) представляет сферу и что оно вообще может представлять? Ответ: кроме сферы оно может представлять точку и пустое множество. (Выясните, когда имеет место каждый из этих случаев.)

Уравнение плоскости.

Теорема 2. *Всякая плоскость представляется в прямоугольных координатах линейным уравнением вида*

$$ax + by + cz + d = 0. \tag{4}$$

Доказательство. Пусть α — данная плоскость и A, B — такие точки, что отрезок AB перпендикулярен плоскости α и делится ею пополам. Плоскость α является геометрическим местом (множеством) точек M, для которых $AM = BM$[12]). Если в равенстве $AM^2 = BM^2$ выписать квадраты расстояний в координатах, то квадраты координат уничтожатся, и мы получим равенство вида (4). Этот вывод совершенно аналогичен соответствующему выводу уравнения прямой на плоскости. □

§ 2. Прямая на плоскости

Прямоугольные координаты на плоскости можно задавать началом координат O и единичными векторами по осям: $\boldsymbol{i}$ по оси x, $\boldsymbol{j}$ — по оси y (рис. 71). Они взаимно ортогональны в силу взаимной перпендикулярности осей. Так что

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = 1, \quad \boldsymbol{ij} = 0. \tag{1}$$

Эти векторы служат основными векторами.

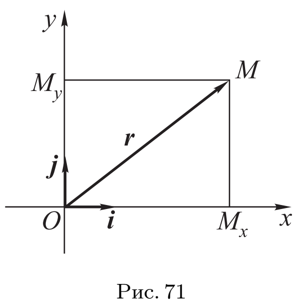


Рис. 71

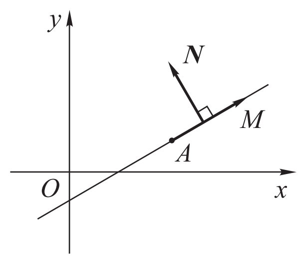


Рис. 72

Вектор $\boldsymbol{r} = \overrightarrow{OM}$, проведенный из начала в точку M, — «радиус-вектор» точки M — представляется в виде:

$$\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}. \tag{2}$$

[12]Это свойство плоскости легко доказывается из соответствующего свойства прямой. Если M — какая-либо точка, то проводим через нее и точки A, B плоскость и в ней применяем теорему о прямой.

Здесь x, y — координаты точки M. По исходному обычному определению это координаты проекций M_x, M_y точки M на оси (рис. 71). Согласно с этим они являются проекциями радиус-вектора $\boldsymbol{r}$ на оси. Скалярно умножая (2) на $\boldsymbol{i}$ и на $\boldsymbol{j}$, найдем (благодаря (1))

$$\overrightarrow{OM}_x = \boldsymbol{ir} = x, \quad \overrightarrow{OM}_y = \boldsymbol{jr} = y. \tag{3}$$

Прямая, очевидно, может быть задана какой-либо точкой A, через которую она проходит, и ненулевым вектором $\boldsymbol{N}$, ей перпендикулярным — «вектором нормали» или «нормалью $\boldsymbol{N}$ к прямой» (рис. 72)[13].

Теорема 1. *Прямая на плоскости представляется линейным уравнением*

$$ax + by + c = 0,$$

причем a, b → это координаты вектора нормали $\boldsymbol{N}$, $-c = \boldsymbol{N} \cdot \overrightarrow{OA}$ — *скалярное произведение вектора нормали на радиус-вектор точки A на прямой.*

Доказательство. Пусть $M(x,\ y)$ — какая-либо точка прямой, заданной точкой A и нормалью $\boldsymbol{N}$. Тогда (если M отлична от A) $\overrightarrow{AM} \perp \boldsymbol{N}$, так что $\overrightarrow{AM} \cdot \boldsymbol{N} = 0$. Если же M совпадает с A, то $\overrightarrow{AM} = \boldsymbol{0}$, поэтому также $\overrightarrow{AM} \cdot \boldsymbol{N} = 0$. С другой стороны, для всякой точки M, не лежащей на данной прямой, вектор $\overrightarrow{AM}$ не ортогонален $\boldsymbol{N}$. Таким образом, точка M лежит на прямой тогда и только тогда, когда $\overrightarrow{AM} \cdot \boldsymbol{N} = 0$. Это равенство представляет поэтому уравнение прямой, стоит лишь развернуть его в координатах.

Пусть x, y — координаты точки M и a, b — координаты нормали $\boldsymbol{N}$, так что $\overrightarrow{OM} = \boldsymbol{i}x + \boldsymbol{j}y$ и $\boldsymbol{N} = \boldsymbol{i}a + \boldsymbol{j}b$. Тогда

$$\overrightarrow{AM} = \overrightarrow{OM} - \overrightarrow{OA} = (\boldsymbol{i}x + \boldsymbol{j}y) - \overrightarrow{OA},$$

и потому из равенства $\overrightarrow{AM} \cdot \boldsymbol{N} = 0$ следует, что

$$\overrightarrow{AM} \cdot \boldsymbol{N} = (\boldsymbol{i}x + \boldsymbol{j}y)(\boldsymbol{i}a + \boldsymbol{j}b) - \overrightarrow{OA} \cdot \boldsymbol{N} = ax + by - \overrightarrow{OA} \cdot \boldsymbol{N} = 0,$$

что и требовалось доказать. $\square$

Таким образом мы не только передоказали еще раз прежнюю теорему 1.2.1 о представлении прямой линейным уравнением, но и выяснили геометрический смысл его коэффициентов.

Нетрудно доказать и обратную теорему.

---

[13] *Нормалью* называется прямая, перпендикулярная прямой (или плоскости).

Теорема 2. *Всякое линейное уравнение представляет прямую, причем коэффициенты при x, y — это координаты некоторой нормали к прямой, а свободный член равен взятому с обратным знаком скалярному произведению этой нормали на радиус-вектор какой-либо точки прямой.*

Д о к а з т е л ь с т в о. Пусть дано уравнение

$$ax + by + c = 0.$$

Берем вектор $\boldsymbol{N}$ с координатами a, b и такую точку A, что $\boldsymbol{N} \cdot \overrightarrow{OA} = -c$. (Такую точку легко найти хотя бы на прямой, идущей через O по вектору $\boldsymbol{N}$. Укажите как.) Через точку A проводим прямую l, перпендикулярную вектору $\boldsymbol{N}$. Тогда по предыдущей теореме уравнение этой прямой и будет данным (рис. 73). □

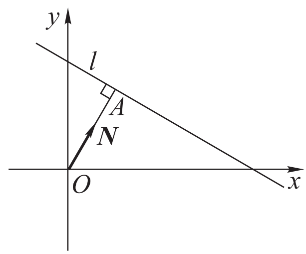


Рис. 73

Вспомним еще теорему: два уравнения задают одну и ту же прямую тогда и только тогда, когда их коэффициенты пропорциональны (дополнение к теореме 1.2.1). Читатель сам ее докажет, заметив в первую очередь, что векторы $\boldsymbol{N}_1$, $\boldsymbol{N}_2$ перпендикулярны одной прямой, тогда и только тогда, когда они параллельны и, стало быть, получаются один из другого умножением на число (ср. доказательство теоремы 3, § 3). □

Важный частный случай уравнения прямой — так называемое *уравнение в нормальной форме*. Это такое уравнение, в котором вектор нормали к прямой единичный, так что $a^2 + b^2 = 1$.

Теорема 3. *Всякая прямая может быть представлена уравнением в нормальной форме:*

$$ax + by = p; \quad a^2 + b^2 = 1, \quad p \geqslant 0. \tag{4}$$

В нем p — расстояние прямой от начала O (т. е. если $p > 0$ — длина опущенного на нее из O перпендикуляра) и a, b, — координаты нормали, направленной от прямой в сторону от начала (или в любую сторону, если прямая проходит через начало).

Д о к а з а т е л ь с т в о. Пусть $\boldsymbol{n}$ — единичный вектор нормали к прямой и a, b — его координаты, так что $a^2 + b^2 = 1$.

По теореме 1 прямая представляется уравнением $ax + by = -c$.

Если же она проходит через начало, то $c = 0$, и уравнение прямой будет

$$ax + by = 0$$

в согласии с утверждением теоремы.

Пусть прямая не проходит через начало, и пусть $\overrightarrow{OP}$ — опущенный из него перпендикуляр — направленный отрезок. Тогда если нормаль направлена по нему, то $\overrightarrow{OP} = p\boldsymbol{n}$, где $p = |\overrightarrow{OP}|$ — длина перпендикуляра — расстояние от O до прямой. Поэтому $\overrightarrow{OP} \cdot \boldsymbol{n} = p > 0$ и нормаль $\boldsymbol{n}$ направлена в сторону от начала, если ее откладывать от точки на прямой (рис. 74, a). Вместе с тем согласно теореме 1

$$-c = \overrightarrow{OP} \cdot \boldsymbol{n} = p.$$

Таким образом, уравнение прямой имеет вид $ax + by = p$, причем a, b, c имеют именно тот смысл, какой указан в теореме. $\square$

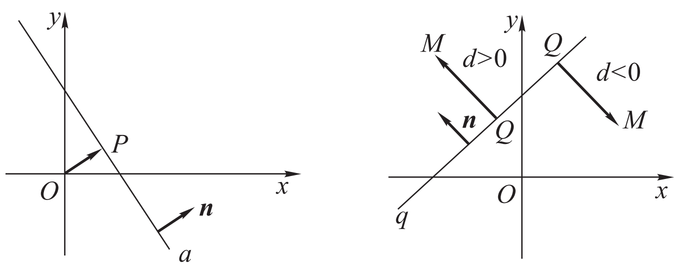


Рис. 74

Следующая теорема устанавливает замечательное свойство уравнения в нормальной форме.

Теорема 4. *Если прямая задана уравнением в нормальной форме, то для всякой точки $M(x, y)$ величина*

$$ax + by - p = \pm d$$

представляет собой расстояние от точки до прямой со знаком плюс, если M лежит от прямой с той стороны, куда направлена нормаль, и со знаком минус, если M лежит с противоположной стороны (*и, значит, с той стороны, где лежит начало O, если прямая через него не проходит*) (рис. 74, $б$).

Доказательство. Пусть прямая q задана уравнением вида (4) — в нормальной форме.

Заметим, что для точек прямой сказанное в теореме верно, так как $d = 0$. Возьмем какую-либо точку $M(x,\ y)$, не лежащую на данной прямой q, и пусть MQ — опущенный из нее перпендикуляр на данную прямую q. Мы имеем

$$\overrightarrow{QM} = \overrightarrow{OM} - \overrightarrow{OQ}, \tag{5}$$

и если $\boldsymbol{n}(a,\ b)$ единичная нормаль, то

$$\boldsymbol{n}\overrightarrow{QM} = \pm|QM| = \pm d. \tag{6}$$

Причем $\boldsymbol{n}\overrightarrow{OM} = +d$, если M лежит с той стороны, куда направлена $\boldsymbol{n}$, и $\boldsymbol{n}\overrightarrow{QM} = -d$ в противоположном случае.

Вместе с тем, по представлению скалярного произведения в координатах,

$$\boldsymbol{n}\overrightarrow{OM} = ax + by, \tag{7}$$

а $\boldsymbol{n}\overrightarrow{OQ} = p$ — расстояние от O до прямой с соответствующим знаком.

Таким образом, из полученных равенств (5), (6), (7)

$$\pm d = \boldsymbol{n}\overrightarrow{QM} = \boldsymbol{n}\overrightarrow{OM} - \boldsymbol{n}\overrightarrow{OQ} = ax + by - p,$$

что и требовалось доказать. □

Координаты единичного вектора — это косинусы углов, которые он образует с положительными направлениями осей. Поэтому если α — угол нормали с положительной полуосью x, то

$$a = \cos\alpha, \quad b = \sin\alpha.$$

Поэтому уравнение в нормальной форме можно написать в виде:

$$x\,\cos\alpha + y\,\sin\alpha = p.$$

Если прямая задана каким-нибудь уравнением

$$ax + by + c = 0,$$

то переход к нормальной форме состоит, во-первых, в переходе к единичной нормали, т. е. к замене a, b и c соответственно на

$$\frac{a}{\pm\sqrt{a^2+b^2}}, \quad \frac{b}{\pm\sqrt{a^2+b^2}}, \quad \frac{c}{\pm\sqrt{a^2+b^2}}.$$

Во-вторых, знак у корня нужно выбрать так, чтобы свободный член получился отрицательным (если только $c \neq 0$), т. е. у корня берется знак, противоположный знаку c.

Если прямая задана общим линейным уравнением $ax + by + c = 0$, то расстояние d можно найти по формуле

$$d = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}}.$$

(Докажите это.)

§ 3. Плоскость и прямая

Общее уравнение плоскости. Выведем уравнение плоскости; этот вывод проводится совершенно так же, как вывод уравнения прямой на плоскости (§ 2).

Пусть дана некоторая плоскость α и выбрано начало отсчета O, из которого проводятся радиус-векторы. Возьмем на плоскости α какую-нибудь точку M_0, и пусть $\boldsymbol{n}$ — какая-нибудь нормаль плоскости α, т. е. перпендикулярный ей ненулевой вектор.

Произвольная точка M лежит на плоскости α и тогда и только тогда, когда вектор $\overrightarrow{M_0M}$ перпендикулярен нормали $\boldsymbol{n}$ или M совпадает с M_0. Это равносильно тому, что равно нулю скалярное произведение векторов $\boldsymbol{n}$ и $\overrightarrow{M_0M}$:

$$\boldsymbol{n} \cdot \overrightarrow{M_0M} = 0.$$

Если $\boldsymbol{r}$, $\boldsymbol{r}_0$ — радиус-векторы точек M, M_0, то $\overrightarrow{M_0M} = \boldsymbol{r} - \boldsymbol{r}_0$. Поэтому написанное равенство равносильно таким:

$$\boldsymbol{n}(\boldsymbol{r} - \boldsymbol{r}_0) = 0; \quad \boldsymbol{nr} = q \quad (q = \boldsymbol{nr}_0). \tag{1}$$

Итак, точка M лежит на данной плоскости в том и только в том случае, когда ее радиус-вектор удовлетворяет уравнению (1), в котором $\boldsymbol{n}$ — нормаль и $\boldsymbol{r}_0$ — радиус-вектор какой-либо точки данной плоскости. Это значит, что (1) есть уравнение плоскости в векторной форме (см. рис. 75, *а*).

Из доказанного выводится

Теорема 1. *В прямоугольных координатах всякая плоскость задается уравнением первой степени.*

Доказательство. По доказанному плоскость задается уравнением (1). Выразим в нем скалярное произведение через координаты.

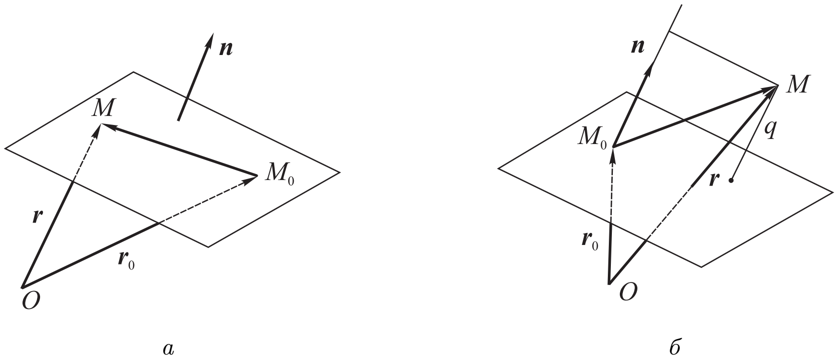


Рис. 75

Координаты данного вектора нормали $\boldsymbol{n}$ обозначим a, b, c. Координаты радиус-вектора $\boldsymbol{r}$ точки M на плоскость — это ее координаты x, y, z. Согласно выражению скалярного произведения в координатах

$$\boldsymbol{nr} = ax + by + cz.$$

Поэтому из второго уравнения (1), полагая $q = -d$, получаем

$$ax + by + cz + d = 0. \tag{2}$$

Это и есть уравнение данной плоскости, поскольку оно — только переписанное в координатах ее векторное уравнение (1). Теорема доказана. □

Перепишем еще в координатах первое из уравнений (1). Если x_0, y_0, z_0 — координаты радиус-вектора $\boldsymbol{r}_0$, то координаты вектора $\boldsymbol{r} - \boldsymbol{r}_0$ суть $x - x_0$, $y - y_0$, $z - z_0$. Поэтому первое уравнение (1) в координатах выглядит так:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \tag{3}$$

Это уравнение плоскости, проходящей через точку $(x_0,\ y_0,\ z_0)$ и имеющей нормаль с координатами a, b, c.

Теорема 2 (обратная теореме 1). *Всякое уравнение первой степени в прямоугольных координатах в пространстве задает плоскость.*

Доказательство. Пусть дано уравнение первой степени

$$ax + by + cz + d = 0. \tag{4}$$

Подберем какие-нибудь x_0, y_0, z_0, удовлетворяющие этому уравнению, так что

$$ax_0 + by_0 + cz_0 + d = 0.$$

Вычитая из (4), получим

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \tag{5}$$

Сравнив это уравнение с уравнением (3), видим, что оно представляет плоскость, проходящую через точку $M_0(x_0,\ y_0,\ z_0)$ и перпендикулярную вектору $\boldsymbol{n} = \{a,\ b,\ c\}$. А уравнение (5) по его выводу равносильно (4)[14]. Следовательно, и уравнение (4) задает плоскость. Теорема доказана. □

Теорема 3. *Два уравнения первой степени задают одну и ту же плоскость в том и только в том случае, когда одно получается из другого умножением на численный множитель.*

Доказательство. Пусть даны два уравнения

$$ax + by + cz + d = 0, \quad a_1x + b_1y + c_1z + d_1 = 0. \tag{6}$$

Если одно получается из другого умножением на некоторое число, то уравнения, очевидно, равносильны и, стало быть, задают одну и ту же плоскость.

Нужно доказать обратное: если уравнения (6) задают одну и ту же плоскость, то одно получается из другого умножением на число.

Допустим, уравнения задают плоскость α. Тогда a, b, c и a_1, b_1, c_1 являются координатами векторов, перпендикулярных α и, стало быть, коллинеарных. Поэтому они пропорциональны:

$$a_1 = \lambda a, \quad b_1 = \lambda b, \quad c_1 = \lambda c. \tag{7}$$

Подставляя (7) во второе уравнение (6), получаем

$$\lambda(ax + by + cz) + d_1 = 0,$$

откуда

$$d_1 = -\lambda(ax + by + cz).$$

---

[14]Уравнения называются *равносильными*, если всякий набор переменных, удовлетворяющий первому, удовлетворяет и второму, и обратно, т. е. в координатах равносильные уравнения это такие, которые задают одну и ту же фигуру — одно и то же множество точек. Это относится к любым уравнениям, не только первой степени.

И так как из первого уравнения (6) $d = -(ax + by + cz)$, то

$$d_1 = \lambda d.$$

Вместе с (7) это и значит, что второе уравнение получается из первого умножением на число, что и требовалось доказать. □

Говорят, что коэффициенты уравнений *пропорциональны*, и пишут:

$$\frac{a_1}{a} = \frac{b_1}{b} = \frac{c_1}{c} = \frac{d_1}{d}.$$

Эта запись и означает, что существует такое число λ, что $a_1 = \lambda a, \ \ldots, \ d_1 = \lambda d$. Она условна, так как не исключает нуля в знаменателе. Но если, например, $c = 0$, то и $c_1 = 0$.

Уравнение плоскости в нормальной форме. Векторным *уравнением плоскости в нормальной форме* называется такое ее уравнение (1), в котором $\boldsymbol{n}$ — единичный вектор нормали и $p \leqslant 0$, т. е. уравнение вида

$$\boldsymbol{nr} = p; \quad |\boldsymbol{n}| = 1, \quad p \geqslant 0. \tag{8}$$

Неравенство $p \geqslant 0$ всегда можно обеспечить, меняя, если нужно, знаки в обеих частях равенства.

В таком случае p — *это расстояние плоскости от начала, а нормаль* $\boldsymbol{n}$ *направлена от плоскости в сторону, противоположную той, где лежит начало* (если же оно лежит на самой плоскости, то направление $\boldsymbol{n}$ безразлично).

Действительно, $\boldsymbol{nr} = p > 0$ тогда и только тогда, когда $\boldsymbol{n}$, $\boldsymbol{r}$ не нулевые и образуют острый угол (рис. 75, a).

Вектор $\boldsymbol{n}$ единичный, поэтому $\boldsymbol{nr}$ — это проекция вектора $\boldsymbol{r}$ на луч, идущий из начала в направлении $\boldsymbol{n}$, т. е. это длина перпендикуляра, опущенного из начала на данную плоскость. А это и есть расстояние от начала до плоскости. □

Расстояние от точки до плоскости. Для уравнения в нормальной форме выполняется следующее красивое соотношение.

Теорема 3. *Для всякой точки M если ее радиус-вектор $\boldsymbol{r}$ подставим в нормальное уравнение плоскости α, то получим*

$$\boldsymbol{nr} - p = \pm q,$$

где q — расстояние от точки M до плоскости α. Знак плюс имеет место, если точка M и начало O лежат с разных сторон от плоскости; если же — с одной стороны, то имеет место минус (на самой плоскости, разумеется, $q = 0$).

Действительно, пусть M — точка, не лежащая на данной плоскости α, а M_0 — точка на α (рис. 75, *б*). Рассмотрим скалярное произведение $\overrightarrow{M_0M} \cdot \boldsymbol{n}$, где $\boldsymbol{n}$ — единичная нормаль к α, направленная, как в нормальном уравнении. Тогда если точка M лежит от плоскости α со стороны, противоположной той, где лежит начало, то $\overrightarrow{M_0M} \cdot \boldsymbol{n} > 0$. Если же — с той же стороны, где начало, то $\overrightarrow{M_0M} \cdot \boldsymbol{n} < 0$.

По абсолютной величине это скалярное произведение есть не что иное, как длина проекции отрезка M_0M на прямую вдоль $\boldsymbol{n}$, т. е. длина перпендикуляра, опущенного из точки M на плоскость α. А это то же, что расстояние q от M до α:

$$|\overrightarrow{M_0M} \cdot \boldsymbol{n}| = q. \tag{9}$$

Теперь заметим, что

$$\overrightarrow{M_0M} = \boldsymbol{r} - \boldsymbol{r}_0,$$

где $\boldsymbol{r}$, $\boldsymbol{r}_0$ — радиус-векторы точек M, M_0. Поэтому

$$\overrightarrow{M_0M} \cdot \boldsymbol{n} = \boldsymbol{n}\boldsymbol{r} - \boldsymbol{n}\boldsymbol{r}_0.$$

А так как точка M_0 лежит на плоскости α, то по уравнению (1)

$$\boldsymbol{n}\boldsymbol{r}_0 = p.$$

Таким образом,

$$\overrightarrow{M_0M} \cdot \boldsymbol{n} = \boldsymbol{n}\boldsymbol{r} - p.$$

Сравнивая с (9) и вспоминая о знаке произведения $\overrightarrow{M_0M} \cdot \boldsymbol{n}$, получаем, что

$$\boldsymbol{n}\boldsymbol{r} - p = \pm q,$$

где q — расстояние от точки M с радиус-вектором $\boldsymbol{r}$ до плоскости α и знаки «—» или «+» берутся в зависимости от того, лежит ли точка M от плоскости α со стороны начала или с противоположной стороны. Что и требовалось доказать. □

Чтобы получить ***нормальное уравнение плоскости*** не в векторах, а ***в прямоугольных координатах,*** достаточно выразить в координатах произведение $\boldsymbol{n}\boldsymbol{r}$. Тогда мы получим

$$ax + by + cz = p,$$

и так как вектор $\boldsymbol{n}$ единичный, то

$$a^2 + b^2 + c^2 = 1. \tag{10}$$

Координаты единичного вектора — это не что иное как его «направляющие косинусы». Если дано уравнение плоскости в общем виде

$$Ax + By + Cz + D = 0, \tag{11}$$

то для того, чтобы получить из него нормальное уравнение, нужно обеспечить равенство (10) и $p \geqslant 0$. Для этого делим на $\pm\sqrt{A^2 + B^2 + C^2}$ и берем знак, противоположный знаку D (чтобы при переносе свободного члена в правую часть там было $p > 0$).

Если же не заботиться о знаке, то можно сформулировать такой результат: расстояние точки $M(x,\ y,\ z)$ до плоскости с уравнением (11) есть

$$q = \left| \frac{Ax + By + Cz + D}{\sqrt{A^2 + B^2 + C^2}} \right|.$$

Замена координат в пространстве. Полученную формулу можно применить к случаю, когда рассматривается одна из координатных плоскостей некоторой прямоугольной системы координат, отличной от исходной. Расстояние до этой плоскости равно (по модулю) соответствующей координате в этой системе. Таким образом, мы получаем важное следствие — то, что переход от одной прямоугольной системы координат в пространстве к другой осуществляется по линейным формулам. (Для плоскости аналогичный факт был доказан в главе I.) Воспользоваться этим нам придется при классификации поверхностей второго порядка в гл. V.

Уравнение плоскости в отрезках на осях. Так называется уравнение вида

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

Во-первых, это — уравнение плоскости, так как оно первой степени. Во-вторых, a, b, c — это координаты тех точек на осях, где их пересекает данная плоскость. Например, при $y = z = 0$, т. е. на оси x, получаем $x = a$, т. е. a — это длина отрезка оси x от начала до точки пересечения, взятая с должным знаком.

Общее уравнение приводится к такому виду, если все коэффициенты не равны нулю: делим на свободный член и меняем знак.

Задачи.

1. Рассмотрите плоскости, в уравнениях которых равны нулю по одному, по два коэффициента и свободный член также может равняться нулю.

2. При каком условии плоскость может быть задана уравнением вида $z = k_1 x + k_2 y + l$? Каков геометрический смысл k_1, k_2, l? Чему равен тангенс угла наклона плоскости к плоскости $z = 0$?

3. Плоскость, проходящая через три точки M_0, M_1, M_2, перпендикулярна $\overrightarrow{M_0M} \times \overrightarrow{M_0M_2}$. Точка M лежит в этой плоскости тогда и только тогда, когда векторы $\overrightarrow{M_0M_1}$, $\overrightarrow{M_0M_2}$, $\overrightarrow{M_0M}$ компланарны. Выразите это условие через смешанное произведение и получите уравнение плоскости в векторной форме, а также в координатах (с определителем).

Задание плоскости двухвекторником. Зададим некоторую плоскость α любой ее точкой $A \in \alpha$ и парой лежащих в α неколлинеарных векторов $\boldsymbol{m}$ и $\boldsymbol{n}$ (рис. 76). Векторы $\boldsymbol{m}$ и $\boldsymbol{n}$ называются *направляющими векторами* плоскости α. Они образуют базис в плоскости α. Любой вектор $\overrightarrow{AX}$, где X — произвольная точка плоскости α, можно разложить по векторам $\boldsymbol{m}$ и $\boldsymbol{n}$ (так как $AX || \alpha$):

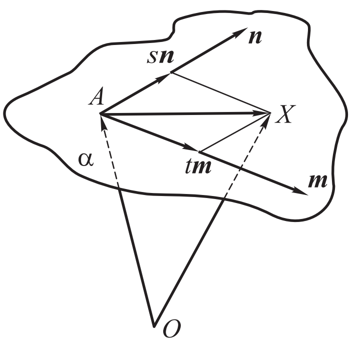


Рис. 76

$$\overrightarrow{AX} = t\boldsymbol{m} + s\boldsymbol{n}. \tag{12}$$

Поскольку $\overrightarrow{OX} = \overrightarrow{OA} + \overrightarrow{AX}$, то, подставляя в это равенство выражение (12) и полагая, как и раньше, $\overrightarrow{OA} = \boldsymbol{r}_0$ и $\overrightarrow{OX} = \boldsymbol{r}$, окончательно получаем

$$\boldsymbol{r} = \boldsymbol{r}_0 + t\boldsymbol{m} + s\boldsymbol{n}. \tag{13}$$

Это уравнение задает плоскость α в пространстве: положение любой точки $X \in \alpha$ определяется заданием упорядоченной пары действительных чисел $(t,\ s)$, причем каждой такой паре соответствует некоторая точка плоскости α. Точка A отвечает пара (0, 0).

Это — *параметрическое* задание плоскости.

§ 4. Прямая в пространстве

Параметрические и канонические уравнения прямой. Всякий вектор, не равный нулю, направленный вдоль прямой, называется *направляющим вектором.*

Прямая l, очевидно, задается любой лежащей на ней точкой M_0 и направляющим вектором $\boldsymbol{v}$. Откладываем от точки M_0 вектор

$\overrightarrow{M_0M_1} = \boldsymbol{v}$ и проводим прямую M_0M_1 (рис. 77).

Точка M лежит на прямой l тогда и только тогда, когда вектор $\overrightarrow{M_0M}$ коллинеарен $\boldsymbol{v}$, т. е. когда существует такое число t, что

$$\overrightarrow{M_0M} = \boldsymbol{v}t. \tag{1}$$

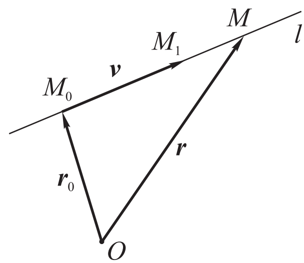


Рис. 77

Если $\overrightarrow{OM_0} = \boldsymbol{r}_0$, $\overrightarrow{OM} = \boldsymbol{r}$ — радиус-векторы точек M_0 и M_1, то $\overrightarrow{M_0M} = \boldsymbol{r} - \boldsymbol{r}_0$, и потому (1) означает

$$\boldsymbol{r} = \boldsymbol{v}t + \boldsymbol{r}_0. \tag{2}$$

При каждом t это дает радиус-вектор некоторой точки M на прямой l; с изменением t «от $-\infty$ до $+\infty$» точка «пробегает» прямую.

Уравнение (2) определяет, таким образом, прямую, задавая каждую ее точку значением переменной t. Эта переменная в такой ее роли называется *параметром*, и уравнение (2) называется (векторным) уравнением прямой в *параметрической форме.*

Его можно переписать в виде трех уравнений для каждой из координат точки на прямой: если a, b, c — координаты вектора $\boldsymbol{v}$, то

$$x = at + x_0, \quad y = bt + y_0, \quad z = ct + z_0. \tag{3}$$

Это равносильно тому, что разности $x - x_0$, $y - y_0$, $z - z_0$ пропорциональны координатам a, b, c:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}. \tag{4}$$

Эти равенства понимаются так: там, где знаменатель — нуль, числитель тоже равен нулю. Равенства (4) с этим условием равносильны равенствам (3). А равенства (3) задают прямую. Стало быть, и равенства (4) определяют прямую. Другими словами: равенства (4) представляют уравнения прямой, проходящей через точку $(x_0,\ y_0\ z_0)$ вдоль вектора $(a,\ b,\ c)$. (Эти уравнения называют *каноническими.*)

Прямая, проходящая через две точки. Если M_1 — еще одна точка прямой, то вектор $\overrightarrow{M_0M_1}$ идет вдоль прямой, и поэтому в уравнении (2) можно взять

$$\boldsymbol{v} = \overrightarrow{M_0M_1} = \boldsymbol{r}_1 - \boldsymbol{r}_0, \tag{5}$$

где $\boldsymbol{r}_1$ — радиус-вектор точки M_1.

Подставляя это в (2), получим

$$\boldsymbol{r} = (\boldsymbol{r}_1 - \boldsymbol{r}_0)t + \boldsymbol{r}_0 = (1-t)\boldsymbol{r}_0 + t\boldsymbol{r}_1. \tag{6}$$

— параметрическое уравнение прямой, проходящей через две данные точки с радиус-векторами $\boldsymbol{r}_1$, $\boldsymbol{r}_0$.

Координаты вектора $\overrightarrow{M_0M_1}$ — это разности координат точек M_1, M_0. Поэтому из (5)

$$a = x_1 - x_0, \quad b = y_1 - y_0, \quad c = z_1 - z_0.$$

Подставляя эти значения в (4), получаем

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = \frac{z - z_0}{z_1 - z_0}. \tag{7}$$

Это, стало быть, уравнения прямой, проходящей через точки $M_0(x_0,\ y_0,\ z_0)$, $M_1(x_1,\ y_1,\ z_1)$.

Прямая как пересечение двух плоскостей. Прямая представляет собою пересечение двух плоскостей, поэтому она задается системой двух уравнений, задающих эти плоскости (рис. 78). Плоскости пересекаются по прямой, если их нормали неколлинеарны. В векторной форме их уравнения будут

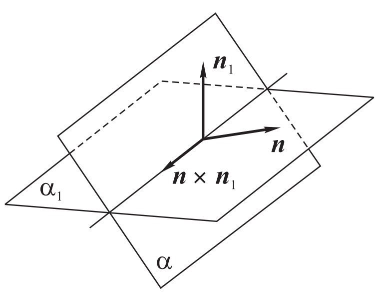


Рис. 78

$$\boldsymbol{n}(\boldsymbol{r} - \boldsymbol{r}_0) = 0, \quad \boldsymbol{n}_1(\boldsymbol{r} - \boldsymbol{r}_1) = 0.$$

Направляющий вектор прямой, по которой они пересекаются, ортогонален их нормалям. В качестве него можно взять векторное произведение $\boldsymbol{n} \times \boldsymbol{n}_1$ (оно отлично от нуля, поскольку $\boldsymbol{n}$, $\boldsymbol{n}_1$ неколлинеарны). Поэтому если $\boldsymbol{r}_0$ — радиус-вектор какой-нибудь общей точки плоскостей α, α_1, то прямую их пересечения можно задать параметрическим уравнением

$$\boldsymbol{r} = (\boldsymbol{n} \times \boldsymbol{n}_1)t + \boldsymbol{r}_0.$$

Задачи.

1. Отсюда получите уравнение в координатах вида (4), где в знаменателях будут координаты векторного произведения.

2. Найти условия параллельности и перпендикулярности плоскостей и прямых, когда они заданы в векторной форме или в координатах уравнениями вида

$$ax + by + cz + d = 0,$$
$$\frac{x - x_0}{k} = \frac{y - y_0}{l} = \frac{z - z_0}{m}.$$

Всего 6 случаев: по два (параллельность и перпендикулярность) для двух плоскостей, для двух прямых, для прямой и плоскости.

§ 5. О задании поверхностей и линий уравнениями

Общие задачи аналитической геометрии в пространстве. Они те же, что на плоскости: представлять фигуры уравнениями и системами уравнений в координатах и из исследования уравнений выводить заключения о геометрических свойствах фигур; выяснять, какие фигуры представляют уравнения того или иного вида. (Помимо уравнений, фигуры могут задаваться также неравенствами.) Исследуемые фигуры — прежде всего те, которые задаются простейшими уравнениями; при этом принято говорить не вообще о фигурах, а о *поверхностях* (а в случае систем двух уравнений — о *линиях* или *кривых*). По числу координат рассматриваются уравнения с тремя переменными x, y, z (т. е. x, y, z — произвольные вещественные числа). При этом *уравнением* с тремя переменными называют соотношение вида $F(x,\ y,\ z) = 0$, где $F(x,\ y,\ z)$ — какая-нибудь функция переменных x, y, z.

Пусть в пространстве дана какая-нибудь поверхность и вместе с тем выбрана некоторая система координат. Как и в случае произвольной фигуры, *уравнением данной поверхности* (в выбранной системе координат) называется такое уравнение с тремя переменными, которому удовлетворяют координаты каждой точки, лежащей на данной поверхности, и не удовлетворяют координаты никакой точки, не лежащей на ней.

Если уравнение дано и мы отвечаем на вопрос, «что такое определяемая им поверхность», то удобно пользоваться следующей формулировкой:

«Поверхность, определяемая данным уравнением (в некоторой си-

стеме координат), есть геометрическое место точек, координаты которых удовлетворяют этому уравнению»[15].

Современной терминологии более соответствует выражение не «геометрическое место», а *множество* точек, что, впрочем, одно и то же. (Употребляется и слово «фигура».) При этом имеется в виду множество **всех** точек, координаты которых удовлетворяют уравнению.

Примеры к этим общим определениям даются уравнениями сферы и плоскости.

В пространственной аналитической геометрии линия, или, что то же, кривая, рассматривается как пересечение двух поверхностей и соответственно определяется системой двух уравнений.

Если $F(x,\ y,\ z) = 0$ и $\Phi(x,\ y,\ z) = 0$ суть уравнения двух поверхностей, пересекающихся по линии L, то линия L есть геометрическое место общих точек этих поверхностей, т. е. точек, координаты которых удовлетворяют одновременно и уравнению $F(x,\ y,\ z) = 0$, и уравнению $\Phi(x,\ y,\ z) = 0$. Говорят: два эти уравнения *совместно определяют линию* L.

Пример представляет задание прямой двумя линейными уравнениями — соответственно тому, что прямая является пересечением двух плоскостей.

Если даны уравнения трех поверхностей, то общие точки этих поверхностей — те, координаты которых удовлетворяют одновременно всем трем уравнениям (и то же для любого числа поверхностей).

Алгебраические уравнения и алгебраические поверхности. Основным предметом изучения в пространственной аналитической геометрии служат поверхности, определяемые в прямоугольных координатах алгебраическими уравнениями. При этом *алгебраическим* называется уравнение $F(x,\ y,\ z) = 0$, в котором левая часть представляет собою многочлен относительно x, y, z с численными коэффициентами, которые, как обычно в алгебре, обозначают в общем виде буквами. Степень многочлена F — наибольшая из степеней составляющих его одночленов — называется *степенью уравнения.* Поверхность, которая в некоторой системе прямоугольных координат определяется алгебраическим уравнением степени n, называется *алгебраической поверхностью порядка* n (или просто поверхностью порядка n). Так,

[15] Данные здесь определения взяты из книги: *Ефимов Н. В.* Краткий курс аналитической геометрии. М.: Гостехиздат, 1954. С. 110–111.

плоскость есть поверхность порядка один (первого порядка), а сфера — поверхность порядка два (второго порядка).

Нетрудно доказать, что если поверхность представляется в одной системе прямоугольных координат уравнением степени n, то она представляется уравнением той же степени и в любой другой системе прямоугольных координат. Следовательно, порядок поверхности не зависит от системы координат, а характеризует саму поверхность.

Вообще свойство, выраженное в прямоугольных координатах, является геометрическим — принадлежит самой фигуре — если оно одинаково имеет место в любой системе прямоугольных координат. (Иначе оно относится к фигуре в связи с данной системой координат.)

Следует заметить, что общая теория алгебраических поверхностей выходит за пределы аналитической геометрии и входит в предмет особой области математики — так называемой *алгебраической геометрии*.

Критика. Все сказанное повторяет применительно к геометрии в пространстве то, о чем говорилось в отношении геометрии на плоскости. И соответственно тому, что там было сказано, сказанное здесь надо понимать с известной долей критики и с условностью. Это относится в первую очередь к основным понятиям уравнения и поверхности.

1. В уравнении $F(x,\ y,\ z) = 0$ левая часть есть функция трех переменных x, y, z и, как «выражение», может не содержать части этих переменных, подобно, например, уравнению $x = 0$. (Впрочем, если написать в виде $x + y + z - y - z = 0$, то стоящее слева выражение будет содержать y и z, — смотря как понимать слово «содержит».)

Поэтому запись уравнения сама по себе его не определяет: должно быть указано, к каким переменным оно относится (как для функции требуется указать область задания): так, например, уравнение можно рассматривать для двух переменных (на плоскости), или для трех (в пространстве), или для одной переменной.

Постоянная — тоже функция: функция, равная одному и тому же числу во всей области задания, в пространстве — для всех x, y, z. Поэтому формально равенства $1 = 0$, $0 = 0$ тоже являются уравнениями. Первое никогда не выполняется и, стало быть, задает пустое множество, второе выполняется при любых x, y, z и, следовательно, определяет все пространство.

Можно исключить эти «уравнения нулевой степени», включив в определение уравнения, что его левая часть $F(x,\ y,\ z)$ есть функция, отличная от постоянной.

Когда уравнение рассматривается как задающее кривую (или по-

верхность), необходимо указывать не только область определения, но и то, к каким координатам оно относится. Например, я спрошу: что задает уравнение $y = 1$? Вы, вероятно, ответите: «прямую» (или «плоскость»), а я скажу: «окружность в полярных координатах, я обозначил радиус игреком».

2. Всякую фигуру, задаваемую системой двух уравнений, можно задать одним уравнением. Действительно, если фигура задается системой

$$\{F(x,\ y,\ z) = 0, \quad G(x,\ y,\ z) = 0\},$$

то она же, очевидно, задается одним уравнением

$$F^2(x,\ y,\ z) + G^2(x,\ y,\ z) = 0.$$

Самый простой пример. Ось z задается как пересечение двух координатных плоскостей $x = 0,\ y = 0$, и она же задается уравнением $x^2 + y^2 = 0$. В этом смысле она является, согласно данному выше общему определению, алгебраической поверхностью второго порядка.

Вообще пересечение любого числа фигур, заданных уравнениями $F_1 = 0,\ \ldots,\ F_n = 0$, можно задать одним уравнением

$$F_1^2 + \ldots + F_n^2 = 0.$$

Точка $(0, 0, 0)$ представляется пересечением трех плоскостей $x = 0$, $y = 0$, $z = 0$ и вместе с тем — одним уравнением

$$x^2 + y^2 + z^2 = 0,$$

так что она тоже поверхность второго порядка.

Вообще термин «поверхность» (как и «кривая») имеет в аналитической геометрии смысл, отличный от того, как его понимают в других частях геометрии в согласии с наглядным представлением.

3. Объединение фигур, представляемых уравнениями $F_1 = 0,\ \ldots,\ F_n = 0$, представляется одним уравнением — их произведением $F_1 \cdot \ldots \cdot F_n = 0$.

Уравнение может представлять пустое множество, как в простейшем случае $x^2 + 1 = 0$. Если уравнение $G = 0$ представляет пустое множество, то

$$FG = 0$$

представляет то же, что $F = 0$.

4. Если пользоваться модулем (как $|x|$ и т. п.), то неравенства можно записывать в виде уравнения. Всякое нестрогое неравенство можно привести к виду $F(x,\ y,\ z) \geqslant 0$, а это равносильно уравнению:

$$F - |F| = 0.$$

(А каким уравнением (с модулем) можно задать строгое неравенство?)

Цилиндры, проекции. В аналитической геометрии *цилиндром* называют фигуру, образованную параллельными прямыми, т. е. являющуюся объединением этих прямых; соответственно эти прямые называются *образующими цилиндра.* Фигура, через каждую точку которой проходит образующая данного цилиндра так, что каждая образующая цилиндра ее пересекает и притом в одной точке, называется *направляющей цилиндра.* Обычно это некоторая кривая, и цилиндр наглядно представляет собою бесконечную поверхность. Пример — круговой цилиндр, направляющей которого служит окружность, а образующие перпендикулярны ее плоскости (рис. 79). Плоскость — тоже цилиндр, направляющей его служит прямая. Особая роль цилиндров в аналитической геометрии состоит в том, что это поверхности, которые задаются уравнениями, не содержащими одной из координат. Это можно выразить утверждением:

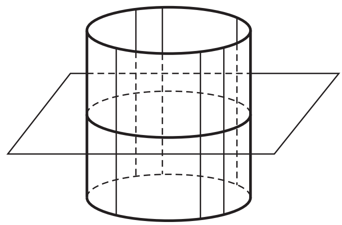


Рис. 79

Теорема. *Уравнение $F(x,\ y) = 0$ в прямоугольных координатах x, y, z задает цилиндр с образующими, параллельными оси z.*

Доказательство. Действительно, при таком уравнении, когда даны x_0, y_0, удовлетворяющие ему, значение координаты z остается ничем не связанным: она может принимать любое значение. Это и значит, что прямая $x = x_0$, $y = y_0$, параллельная оси z, содержится в фигуре с данным уравнением, и так как это верно для любых $(x_0,\ y_0)$), удовлетворяющих уравнению, то вся фигура покрыта такими прямыми. То есть она представляет собой цилиндр с образующими, параллельными оси z. $\square$

С другой стороны, если имеется цилиндр, то выберем прямоугольные координаты так, чтобы ось z была параллельна его образующим. Образующие пересекают тогда плоскость xy, и точки пересечения образуют фигуру — направляющую цилиндра. Представляя ее уравнением $F(x,\ y) = 0$, получим тем самым уравнение цилиндра.

Если поверхность, заданную уравнением $F(x,\ y,\ z) = 0$, мы пересечем плоскостью $z = c = \text{const}$, то получим кривую L, задаваемую системой двух уравнений: $\{F(x,\ y,\ c) = 0,\ z - c = 0\}$. В первом уравнении только две переменных, и оно задает поэтому в пространстве цилиндр с образующими, параллельными оси z, а на плоскости xy — кривую с тем же уравнением $F(x,\ y,\ c) = 0$. Эта кривая является проекцией кривой L.

Такими соображениями постоянно пользуются, например, изучая форму поверхности по ее сечениям, когда в сечениях получаются уже известные кривые.

Цилиндры представляют собой частный случай поверхностей, образуемых (иногда говорят порождаемых) прямыми; эти прямые называют прямолинейными образующими или, короче, образующими поверхности, а сами поверхности называют линейчатыми. Примером может служить конус, как его понимают в аналитической геометрии, — поверхность, образуемая прямыми, проходящими через одну точку.

Глава V

ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА

§ 1. Разные типы поверхностей второго порядка

Определение. Подобно кривой второго порядка, *поверхностью второго порядка* — сокращенно **ПВП** — называется множество точек, прямоугольные координаты которых удовлетворяют уравнению второй степени:

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{23}yz + \\ + 2a_{13}xz + 2a_1x + 2a_2y + 2a_3z + a_0 = 0. \quad (1)$$

Предполагается, что хотя бы один из коэффициентов при членах второй степени не равен нулю (иначе уравнение было бы первой степени).

Это свойство сохраняется при преобразовании к любым другим прямоугольным координатам. Это доказывается буквально так же, как в случае КВП (т. е. в случае уравнения с двумя переменными).

При преобразовании координат координаты x, y, z выражаются через «новые» координаты x', y', z' линейно. Поэтому степень уравнения не может повышаться. Но и понизиться она не может, так как иначе

при обратном преобразовании не получалось бы уравнение второй степени,

Задача, как и в случае КВП, состоит в том, чтобы выяснить, какие есть типы ПВП, какого геометрического вида. Но прежде чем перечислить их, установим одно их общее свойство.

Лемма. *Пересечение* ПВП *с плоскостью представляет собой* КВП, *кроме того особого случая, когда оно оказывается всей плоскостью.* (Подобно тому, как КВП может представлять собой прямую или две прямые, ПВП может быть плоскостью или парой плоскостей, задаваемых, например, уравнением $z^2 = a^2$ (так что $z = \pm a$ и $z = 0$, если $a = 0$).)

Доказательство. Пусть даны какая-то ПВП и некоторая плоскость α. Можно преобразовать координаты так, чтобы она стала плоскостью $z = 0$. Уравнение данной ПВП будет иметь тот же вид (1). Уравнение сечения этой ПВП плоскостью α получается из него, если в нем положить $z = 0$, так что получится уравнение без z. Для этого уравнения есть разные возможности.

а) Оно содержит члены второй степени и, стало быть, представляет КВП.

Однако члены второй степени могут исчезнуть (если все они содержали z). Тогда может быть еще три случая:

б) Получается линейное уравнение: оно представляет прямую, а прямая тоже есть КВП.

в) Линейных членов тоже не будет (если они содержали только z), и останется только свободный член: уравнение будет $a_0 = 0$; и если $a_0 \neq 0$, оно представляет пустое множество, которое тоже есть КВП.

г) Наконец, может случиться, что все члены исчезнут и уравнение сведется к $0 = 0$. Оно всегда выполняется, ему удовлетворяют любые x, y, так что оно представляет всю плоскость. □

Этот разбор разных случаев поучителен, он показывает, как за деталями могут обнаруживаться возможные существенные особенности рассматриваемого вопроса.

Оказывается, *есть* 15 *разных типов* ПВП.

Перечислим их, указав уравнения, которыми они задаются в подходящих координатах. Эти уравнения называются *каноническими.*

1. ***Эллипсоид*** (рис. 80)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{2}$$

В частности, при $a = b = c$ получаем сферу

$$x^2 + y^2 + z^2 = a^2.$$

Из уравнения (2) очевидно, что

$$|x| \leqslant a, \quad |y| \leqslant b, \quad |z| \leqslant c. \tag{3}$$

Это значит, что эллипсоид содержится в прямоугольном параллелепипеде, задаваемом неравенствами (3) (рис. 81).

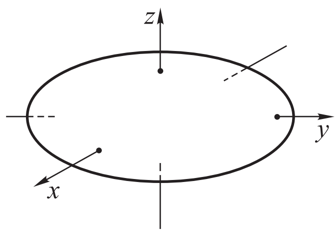


Рис. 80

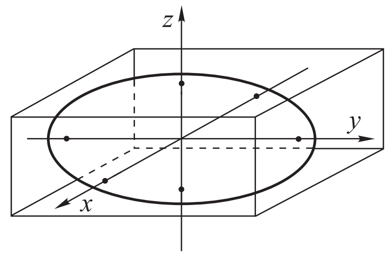


Рис. 81

Сечение эллипсоида плоскостью $z = 0$ представляет эллипс

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

То же верно для сечений плоскостями $x = 0$, $y = 0$. Пересечением эллипсоида с плоскостью, если он вообще имеет с нею общие точки, будет либо эллипс, либо одна точка. В самом деле, эллипсоид ограничен. Поэтому по доказанной лемме это пересечение может быть только ограниченной КВП, т. е. либо эллипсом, либо одной точкой. В последнем случае плоскость — касательная к эллипсоиду (рис. 82).

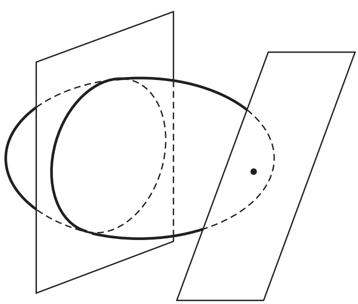


Рис. 82

Далее идут два типа ПВП, называемых *гиперболоидами*.

2. ***Однополостный гиперболоид*** (рис. 83)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \tag{4}$$

3. ***Двуполостный гиперболоид*** (рис. 84)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1 \tag{5}$$

(числа a, b, c в уравнениях (2), (4), (5) называют *полуосями* эллипсоида и гиперболоидов).

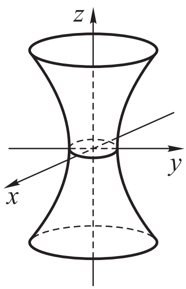


Рис. 83

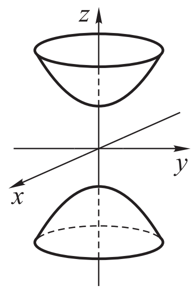


Рис. 84

Плоскость $z = 0$ пересекает первый гиперболоид по эллипсу $x^2/a^2 + y^2/b^2 = 1$.

Это сечение называется *горловым*; от него гиперболоид расширяется в обе стороны.

У двуполостного гиперболоида иначе: никакая плоскость $z = z_0$ при $|z_0| < c$ его вовсе не пересекает. Действительно, при $z = z_0$ уравнение (5) дает

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z_0^2}{c^2} - 1 < 0.$$

А это уравнение определяет пустое множество.

Следовательно, двуполостный гиперболоид состоит из двух «полостей», разделенных целым слоем между двумя параллельными плоскостями $z = \pm c$.

Плоскость $y = 0$ пересекает гиперболоиды по гиперболам (рис. 85, *а*, *б*)

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1, \quad \frac{x^2}{a^2} - \frac{z^2}{c^2} = -1.$$

Аналогично, плоскость $x = 0$ пересекает гиперболоиды по гиперболам. Впрочем, всякая плоскость, проходящая через ось z, пересекает гиперболоиды по гиперболам. Действительно, по доказанной лемме

сечения представляют собою КВП. В данном случае непосредственно видно, что эти кривые состоят из двух бесконечных ветвей (не прямых). А из всех КВП так устроены только гиперболы.

4. ***Конус*** (рис. 86)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

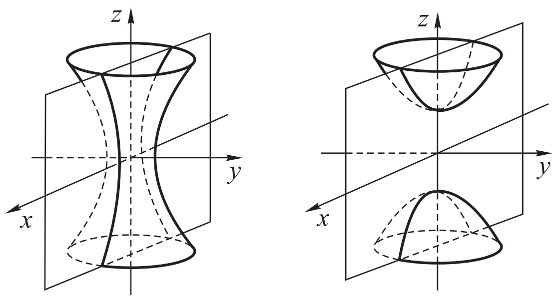


Рис. 85

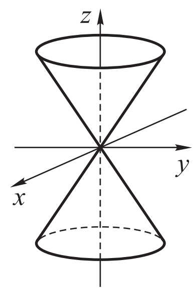


Рис. 86

Эта поверхность состоит из прямых, пересекающихся в одной точке — *вершине* конуса (при выбранных координатах, когда уравнение имеет написанный вид, вершина — в начале координат).

Действительно, если x_0, y_0, z_0 удовлетворяют данному уравнению, то ему удовлетворяют также $x = x_0 t$, $y = y_0 t$, $z = z_0 t$ при любом t. Эти равенства с параметром t представляют прямую (если не все x_0, y_0, z_0 равны нулю), проходящую через начало и точку $(x_0,\ y_0,\ z_0)$. Значит, вместе с любой своей точкой $(x_0,\ y_0,\ z_0)$ конус содержит и всю такую прямую. Он состоит из таких прямых — *образующих*.

Поверхности всех четырех указанных типов имеют по три плоскости симметрии — это координатные плоскости (так как в уравнение входят только x^2, y^2, z^2, и поэтому при перемене знака любой из координат уравнение не изменяется — например, при перемене знака z, так что поверхность симметрична относительно плоскости $z = 0$).

Все эти поверхности имеют центр симметрии: их уравнения не изменяются при перемене знака всех трех координат, так что поверхности симметричны относительно начала координат.

Рассмотрим еще некоторые сечения всех четырех поверхностей.

Если в уравнении любой из них взять $z = \mathrm{const}$, то получим урав-

нение вида
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = p = \mathrm{const}.$$

Если $p > 0$, то это уравнение представляет эллипс на плоскости xy. Это означает, что каждая из рассматриваемых поверхностей пересекается плоскостью $z = \mathrm{const}$ по эллипсу, если плоскость ее действительно «пересекает». При $p < 0$ пересечение пусто; при $p = 0$ оно представляет собой одну точку (с $x = y = 0$), в этом случае плоскость — касательная.

Все эти эллипсы подобны, так как их полуоси равны $a\sqrt{p}$, $b\sqrt{p}$, и, значит, их отношение не зависит от p.

Если $a = b$, то эти эллипсы представляют собою окружности и, стало быть, рассматриваемые поверхности — всех четырех типов — представляют собой *поверхности вращения.* Эллипсоид получается вращением эллипса вокруг оси, гиперболоиды — вращением гипербол, конус — вращением прямой.

Сечения конуса плоскостями, не проходящими через вершину, представляют, согласно лемме, КВП. Если плоскость пересекает одну половину конуса, и притом все его образующие, то сечение ограничено и оно — эллипс; если плоскость параллельна одной из образующих, то сечение состоит из одной бесконечной линии и, стало быть, является параболой. Если же плоскость пересекает обе половины конуса, то сечение — гипербола (рис. 23).

Далее идут **поверхности, канонические уравнения которых не содержат z^2**. Это прежде всего **параболоиды**; их два типа.

5. ***Эллиптический параболоид*** (рис. 87)
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2z.$$

6. ***Гиперболический параболоид*** (рис. 88)
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2z.$$

Сечения этих поверхностей плоскостью $y = 0$ (как и $x = 0$) представляют собой параболы
$$x^2 = 2a^2 z \quad (y^2 = 2b^2 z \quad \text{и} \quad y^2 = -2b^2 z).$$

Сечения же плоскостями $z = p = \mathrm{const} > 0$ представляют собой у эллиптического параболоида эллипсы
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2p.$$

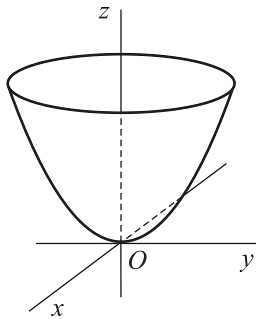


Рис. 87

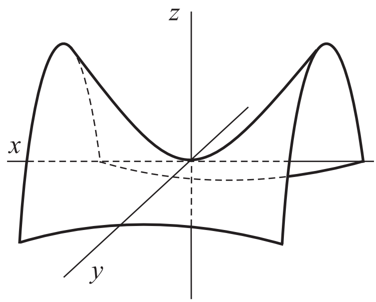


Рис. 88

В частности, это окружности, если $a = b$; тогда это — параболоид вращения.

У гиперболического параболоида сечения плоскостями $z = p = \text{const} \neq 0$ представляют гобой гиперболы

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2p.$$

Сечение плоскостью $z = 0$ представляет собой пару прямых.

Параболоиды имеют по две плоскости симметрии: $x = 0$ и $y = 0$, но не $z = 0$; и они не имеют центра симметрии.

Цилиндры. Далее следуют поверхности, представляющие собою цилиндры, направляющими которых служат кривые второго порядка. Такой цилиндр получается, если через все точки данной КВП провести прямые, перпендикулярные плоскости, содержащей эту КВП. Так как типов КВП всего 8, то и цилиндров, на них построенных, 8 типов. Особый случай представляет КВП, являющаяся пустым множеством: через нее не проходит никакая прямая. Этому соответствует ПВП, являющаяся пустым множеством, Без этого особого случая имеется 7 видов КВП и соответственно 7 видов цилиндрических ПВП. Вот три вида цилиндров на конических сечениях.

7. ***Эллиптический цилиндр*** (рис. 89)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

8. ***Гиперболический цилиндр*** (рис. 90)

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

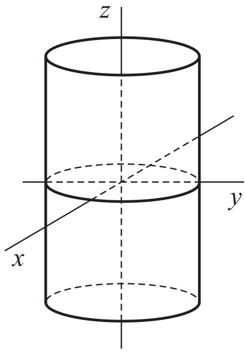


Рис. 89

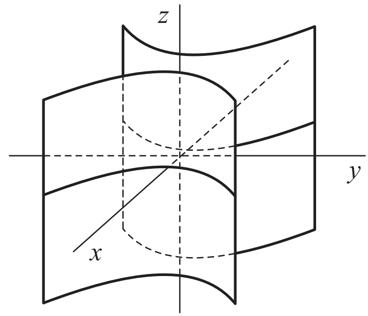


Рис. 90

9. ***Параболический цилиндр*** (рис. 91)

$$y^2 = 2px.$$

За ними идут **ПВП, состоящие из плоскостей** (это тоже цилиндры, поскольку плоскость «состоит» из параллельных прямых).

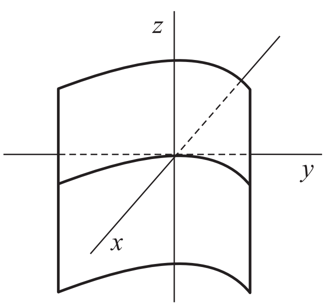


Рис. 91

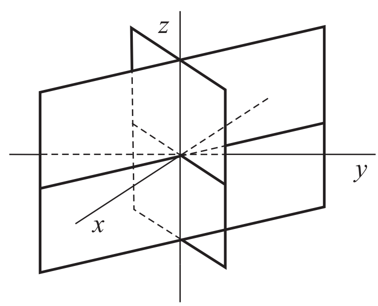


Рис. 92

10. ***Пара пересекающихся плоскостей*** (рис. 92)

$$y^2 - k^2x^2 = 0, \quad k \neq 0.$$

11. ***Пара параллельных плоскостей*** (рис. 93)

$$y^2 - k^2 = 0.$$

12. ***Плоскость*** (рис. 94)

$$y^2 = 0.$$

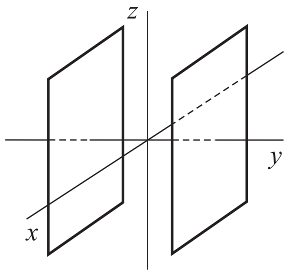


Рис. 93

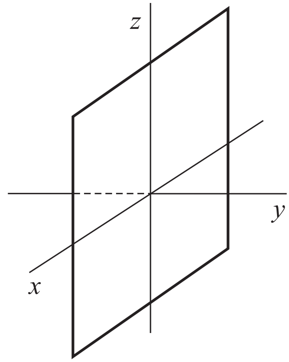


Рис. 94

13. ***Прямая*** — «цилиндр», построенный на одной точке (рис. 95)

$$x^2 + y^2 = 0.$$

Остаются еще два особых случая,
14. ***Одна точка*** (рис. 96)

$$x^2 + y^2 + z^2 = 0$$

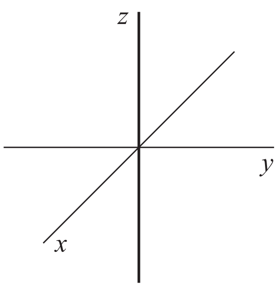


Рис. 95

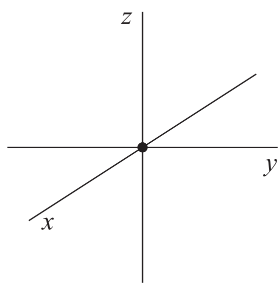


Рис. 96

15. ***Пустое множество*** (рис. 97); его представляют три типа канонических уравнений

$$x^2 + y^2 + z^2 = -1, \quad x^2 + y^2 = -1, \quad x^2 = -1.$$

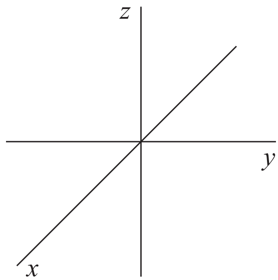


Рис. 97

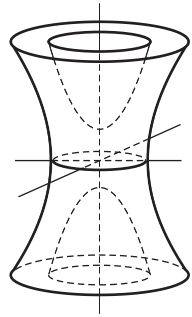


Рис. 98

Сопряженные гиперболоиды. Асимптотический конус. Рассмотрим однополостный и двуполостный гиперболоиды, представленные уравнениями

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1,$$

с одинаковыми полуосями a, b, c (рис. 98).

Сечения их плоскостью $y = 0$ представляют собой *сопряженные* гиперболы

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1, \quad \frac{x^2}{a^2} - \frac{z^2}{c^2} = -1.$$

Аналогичное верно для сечений плоскостью $x = 0$ и вообще для сечений плоскостями, проходящими через ось z, как в этом можно убедиться. Соответственно и сами гиперболоиды называются *сопряженными.*

Рассмотрим вместе с ними конус K, представляемый уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

Покажем, что *при удалении точки по гиперболоиду* (так что абсолютная величина ее координаты z неограниченно возрастает) *точка эта неограниченно приближается к конусу K.*

Пусть $M(x,\ y,\ z)$ — точка на гиперболоиде, а $N(x,\ y,\ z')$ — точка на конусе K с теми же x, y (т. е. с той же проекцией на плоскость $z = 0$). Пусть, для простоты, $z,\ z' > 0$. Расстояние $MN = |z - z'|$.

Вычитая из левых частей уравнений гиперболоидов левую часть уравнения конуса, найдем для первого гиперболоида

$$\frac{(z')^2}{c^2} - \frac{z^2}{c^2} = 1, \quad |z' - z|(z' + z) = c^2,$$

и для второго точно так же $|z' - z|(z' + z) = c^2$.

Таким образом, для точек на обоих гиперболоидах расстояние их от соответствующей точки конуса будет

$$MN = |z' - z| = \frac{c^2}{z' + z}.$$

Поэтому при $(z' + z) \to \infty$ получается $|z' - z| \to 0$, т. е. точки M и N неограниченно сближаются. $\square$

Это свойство выражают, говоря, что конус K является ***асимптотическим конусом*** обоих гиперболоидов.

Сечения этих гиперболоидов плоскостями, проходящими через ось z — ось гиперболоидов, — представляют собою гиперболы, а образующие конуса K являются их асимптотами.

Будем изменять полуоси гиперболоидов в одно и то же число раз, заменяя a на pa и т. д. Так что уравнения получающихся гиперболоидов можно записать так:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = p^2, \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -p^2.$$

При $p \to 0$ эти уравнения «переходят» в уравнение конуса K. Это значит, что при пропорциональном неограниченном уменьшении полуосей оба гиперболоида стягиваются к конусу. Можно сказать, что при изменении параметра p, когда он проходит через нуль, один гиперболоид «превращается» в другой, проходя через их общий асимптотический конус.

Пропорциональное изменение полуосей есть нечто иное, как гомотетия с центром в центре гиперболоидов — начале координат. Бесконечное гомотетичное сжатие гиперболоидов стягивает их к асимптотическому конусу.

§ 2. Приведение квадратичной формы к каноническому виду

Нам надо доказать, что все ПВП исчерпываются пятнадцатью их типами, рассмотренными в § 1. Для этого нам понадобится один важный результат из алгебры.

Определение. *Квадратичной формой* (двух, трех и т. д. переменных) называют однородный многочлен второй степени. Например:

$$2x^2 + 3xy + 4y^2 + z^2$$

— квадратичная форма от трех переменных[16]. Форма имеет «канонический вид» или является «суммой квадратов», если в ней нет членов с произведениями переменных, а только члены с их квадратами (при них возможны и отрицательные коэффициенты).

Удобно представлять переменные как прямоугольные координаты в пространстве, и, соответственно, форма представляется как заданная на пространстве (квадратичная) функция точки $f(X)$. Координаты можно преобразовывать (не сдвигая начала), и форма будет изменяться по своему виду.

Теорема 1. *Всякую квадратичную форму трех переменных можно превратить в «сумму квадратов» путем подходящего преобразования переменных, соответствующего переходу от одних прямоугольных координат к другим.*

Доказательство. Рассмотрим квадратичную форму F трех переменных x, y, z. Представим переменные как прямоугольные координаты в пространстве, так что форма F будет заданной на пространстве функцией точки $f(X)$. Рассмотрим ее на единичной сфере S, т. е. на сфере радиуса единица, с уравнением

$$x^2 + y^2 + z^2 = 1. \tag{1}$$

С алгебраической точки зрения это значит, что мы связываем переменные x, y, z условием (1).

Итак, $f(X)$ будет обозначать данную квадратичную форму как функцию на сфере S. Эта функция непрерывная и поэтому достигает на сфере S своего наибольшего значения; обозначим его a. Оно достигается в некоторой точке A.

---

[16] Вообще, *формой* называют однородный многочлен от нескольких переменных, степени n, если составляющие его одночлены имеют вид $ax_1^{k_1} \ldots x^{k_m}$ с $k_1 + \ldots + k_m = n$.

Повернем оси координат так, чтобы ось x прошла через точку A. Докажем, что тогда в этих новых координатах, которые мы обозначим также x, y, z, форма F представляется в виде

$$f = ax^2 + G(y,\ z), \tag{2}$$

где $G(y,\ z)$ — квадратичная форма только двух переменных, т. е. в формуле (2) члены с произведениями xy, xz отсутствуют, а коэффициент при x^2 — это, как было обозначено, наибольшее значение формы на сфере S, т. е. при условии (1).

Допустим, форма в новых координатах имеет общий вид

$$a_{11}x^2 + 2a_{12}xy + \ldots$$

По выбору координат наибольшее значение (при условии (1)) достигается на оси x, т. е. при $x = 1$, $y = z = 0$. А это значение будет a_{11}. Значит, $a_{11} = a$, как и записано в формуле (2).

Допустим, вопреки тому, что нужно доказать, что $a_{12} \neq 0$ (если $a_{12} = 0$, то рассмотрим a_{13}). Положим $z = 0$. Форма примет вид (так как $a_{11} = a$)

$$f = ax^2 + 2a_{12}xy + a_{22}y^2. \tag{3}$$

Мы рассматриваем ее при условии (1), т. е. (так как $z = 0$) при

$$x^2 + y^2 = 1, \quad x = \sqrt{1 - y^2}.$$

Поэтому из (3) получаем

$$f = a(1 - y^2) + 2a_{12}y\sqrt{1 - y^2} + a_{22}y^2. \tag{4}$$

Возьмем y того же знака, что a_{12}; тогда

$$a_{12}y > 0.$$

Заметим еще, что при $|y| < 1$ $\sqrt{1 - y^2} > 1 - |y|$. Поэтому заменяя в (4) $\sqrt{1 - y^2}$ на меньшую величину и объединяя все члены с y^2, можно написать:

$$f > a + 2a_{12}y + cy^2.$$

Если y достаточно мало, то величина $|cy^2|$ будет мала в сравнении с $a_{12}y$, а поэтому окажется

$$2a_{12}y + cy^2 > 0,$$

и в результате

$$f > a.$$

Но это противоречит тому, что a — это наибольшее значение функции f при условии (2). Следовательно, не может быть $a_{12} \neq 0$, т. е. $a_{12} = 0$.

Совершенно так же должно быть и $a_{13} = 0$.

Таким образом, форма имеет вид (2).

Теперь достаточно привести форму $G(y,\ z)$ к сумме квадратов поворотом осей y, z, и мы приведем исходную форму F к сумме квадратов. Теорема доказана. □

§ 3. Классификация ПВП

Выше в § 1 были перечислены 15 типов ПВП. Докажем теперь, что других ПВП не бывает.

Теорема 1. *Каждая* ПВП *принадлежит к одному из* 15 *типов, перечисленных в* § 1.

Доказательство проводится совершенно подобно тому, как в § 7 гл. II было доказано, что есть только 8 типов КВП.

Пусть дано какое-либо уравнение второго порядка с тремя переменными, которые понимаем как прямоугольные координаты в пространстве. Постараемся преобразовать координаты так, чтобы уравнение приняло канонический вид — один из тех, что указан в § 1.

Прежде всего, согласно доказанной в § 2 теореме, можно выбрать координаты так, чтобы квадратичная форма, которую образуют члены второй степени, привелась к сумме квадратов. Поэтому уравнение в этих координатах x, y, z будет иметь вид

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_1x + 2a_2y + 2a_3z + a_0 = 0. \tag{1}$$

Дальше будем подбирать перенос начала координат. Докажем, что если $a_{11} \neq 0$, то можно перенести начало, сделав преобразование

$$x = x' + p, \tag{2}$$

так что член с x исчезнет (если его нет, то и перенос равен нулю: $p = 0$). Аналогичное, конечно, верно для y и z.

Подставляя в уравнение (1) выражение (2), получим, обозначая члены без x многоточием:

$$a_{11}(x'+p)^2 + 2a_1(x'+p) + \ldots = \\ = a_{11}x'^2 + 2a_{11}x'p + a_{11}p^2 + 2a_1x' + 2a_1p + \ldots$$

Если $a_{11} \neq 0$, то можно взять

$$p = -\frac{a_1}{a_{11}}.$$

Тогда $a_{11}p + a_1 = 0$, так что члены с x' исчезают и остается только $a_{11}(x')^2$.

Теперь мы различаем три возможных случая:

I. В (1) есть все три квадрата.

II. В (1) только два квадрата.

III. В (1) только один квадрат.

Случай I. В этом случае, согласно доказанному, можно, перенося начало, исключить все первые степени переменных. Уравнение (1) приведется к виду (для простоты обозначаем новые переменные x' как старые и новый свободный член обозначаем так же a_0)

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + a_0 = 0. \tag{3}$$

(I, 1) Пусть $a_0 \neq 0$. Тогда, разделив на $-a_0$, можно переписать уравнение в виде

$$\overline{a}_{11}x^2 + \overline{a}_{22}y^2 + \overline{a}_{33}z^2 = 1. \tag{4}$$

(I, 1А) Пусть все коэффициенты положительны. Тогда (4) можно переписать (полагая $\overline{a}_{11} = a^{-2}$ и т. д.) в виде

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Уравнение представляет эллипсоид.

(I, 1Б) Пусть один коэффициент в (4) отрицателен: скажем, $\overline{a}_{33} < 0$. Тогда, полагая $\overline{a}_{11} = a^{-2}$, $\overline{a}_{22} = b^{-2}$, $\overline{a}_{33} = -c^{-2}$, получим

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

Уравнение представляет однополостный гиперболоид.

(I, 1В) Пусть два коэффициента в (4) отрицательны: скажем, $\overline{a}_{11} < 0$, $\overline{a}_{22} < 0$, $\overline{a}_{33} > 0$. Тогда, полагая

$$\overline{a}_{11} = -a^{-2}, \quad \overline{a}_{22} = -b^{-2}, \quad \overline{a}_{33} = c^{-2},$$

получим

$$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad \text{т. е.} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \tag{5}$$

Это — уравнение двуполостного гиперболоида. (Понятно, что знаки коэффициентов могут распределиться иначе, но тогда переименуем координаты так, чтобы получались нужные уравнения (5).)

(I, 1Г) Пусть все три коэффициента в (4) отрицательны. Тогда уравнение (4) не может выполняться ни при каких x, y, z. Оно, стало быть, представляет пустое множество.

(I, 2) Теперь рассмотрим случай, когда в (3) $a_0 = 0$, так что уравнение приобретает вид

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 = 0. \tag{6}$$

(I, 2А) Если здесь все коэффициенты одного знака, то уравнение выполняется только при $x = y = z = 0$, т. е. оно представляет точку.

(I, 2Б) Пусть первые два коэффициента — одного знака, а третий — другого. Можно считать, что два коэффициента положительные, а один — отрицательный (иначе меняем в левой части знак). Тогда положим $a_{11} = a^{-2}$, $a_{22} = b^{-2}$, $a_{33} = -c^{-2}$ и получим

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

Это — уравнение конуса.

Случай II. Теперь рассмотрим второй случай, — когда в (1) один коэффициент при квадратах равен нулю. Можно считать, что $a_{33} = 0$, но $a_{11} \neq 0$, $a_{22} \neq 0$. Тогда члены первой степени с x и y можно исключить, но z может остаться.

(II, 1) Пусть уравнение имеет вид

$$a_{11}x^2 + a_{22}y^2 + 2a_3z + a_0 = 0, \quad a_3 \neq 0.$$

Тогда полагаем

$$z = z' - \frac{a_0}{2a_3}, \quad 2a_3z + a_0 = 2a_3z'$$

и делим на $-a_3$. Уравнение примет вид (обозначаем z' через z)

$$\overline{a}_{11}x^2 + \overline{a}_{22}y^2 = 2z.$$

(II. 1А) Если $\overline{a}_{11},\ \overline{a}_{22} > 0$, то полагаем $\overline{a}_{11} = a^{-2}$, $\overline{a}_{22} = b^{-2}$ и получаем

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2z.$$

Это — уравнение эллиптического параболоида.

Если $\overline{a}_{11}, \overline{a}_{22} < 0$, то заменяя z на $-z$ (переворачивая ось z), получим $\overline{a}_{11},\ \overline{a}_{22} > 0$, т. е. тот же результат.

(II, 1Б) Если же $\overline{a}_{11}, \overline{a}_{22}$ разных знаков, то можно считать $\overline{a}_{11} > 0$, $\overline{a}_{22} < 0$ и переписать уравнение в виде

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2z.$$

Это — уравнение гиперболического параболоида.

(II, 2) Теперь рассмотрим случай, когда уравнение (3) имеет вид

$$a_{11}x^2 + a_{22}y^2 + a_0 = 0.$$

Тут можно действовать совершенно так же, как в случае, когда уравнение имело вид (3).

(II, 2А) Если $a_0 \neq 0$, то, деля на $-a_0$, получаем

$$\overline{a}_{11}x^2 + \overline{a}_{22}y^2 = 1.$$

(а) Если $\overline{a}_{11}, \overline{a}_{22} > 0$, то уравнение приводится к виду

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Это — эллиптический цилиндр.

(б) Если $\overline{a}_{11} > 0$, $\overline{a}_{22} < 0$, то уравнение приводится к виду

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Это — гиперболический цилиндр. (Если же $\overline{a}_{11} < 0$, $\overline{a}_{22} > 0$, то переименовываем координаты.)

(в) $\overline{a}_{11} < 0$, $\overline{a}_{22} < 0$, тогда уравнение не удовлетворяется ни при каких x, y и представляет пустое множество.

(II, 2Б) Если $a_0 = 0$, то уравнение имеет вид

$$a_{11}x^2 + a_{22}y^2 = 0.$$

(а) Если a_{11}, a_{22} разных знаков, то уравнение можно написать в виде

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0.$$

Это — пара пересекающихся плоскостей: $ay = \pm bx$.

(б) Если a_{11}, a_{22} одного знака, то уравнение выполняется только при $x = y = 0$. Это — прямая: ось z.

Случай III. Рассмотрим третий случай, когда в уравнении (1) только один квадрат, например y^2; тогда можно исключить член с y в первой степени и получить уравнение

$$a_{22}y^2 + 2a_1x + 2a_3z + a_0 = 0. \tag{7}$$

(III, 1) Пусть $a_3 \neq 0$ (или $a_1 \neq 0$). Тогда можно исключить a_0, как выше в (II, 1). Если $a_1 \neq 0$ и $a_3 \neq 0$, то повернем оси x, z, т. е. произведем преобразование вида (выпишем только $\overline{x}$)

$$\overline{x} = x\cos\alpha - z\sin\alpha,$$

где

$$\cos\alpha = \frac{a_1}{\sqrt{a_1^2 + a_3^2}}, \quad \sin\alpha = \frac{-a_3}{\sqrt{a_1^2 + a_3^2}},$$

так что

$$a_1x + a_3z = \sqrt{a_1^2 + a_3^2}\,\overline{x}.$$

После этих двух преобразований и деления на a_{22} уравнение примет вид

$$y^2 - 2px = 0, \quad y^2 = 2px.$$

Это — параболический цилиндр.

(III, 2) Пусть в уравнении (7) нет членов первой степени. Тогда оно имеет вид

$$a_{11}y^2 + a_0 = 0.$$

(III, 2A) Если $a_0 \neq 0$, $a_{11} \neq 0$ и a_0, a_{11} разных знаков, то

$$y^2 = -\frac{a_{11}}{a_0} > 0.$$

Это — пара параллельных плоскостей: $y = \pm\sqrt{-a_{11}/a_0}$.

(III, 2Б) Пусть $a_0 \neq 0$, $a_{11} \neq 0$, но a_0, a_{11} одного знака. Тогда уравнение не имеет решений и представляет пустое множество.

(III, 2B) Пусть $a_0 = 0$. Уравнение будет $y^2 = 0$. Оно представляет одну плоскость — плоскость xz.

§ 4. Прямолинейные образующие ПВП

Прямолинейные образующие есть у конусов и цилиндров, но они есть еще у ПВП двух других типов.

Прямолинейные образующие гиперболического параболоида.

Теорема 1. *У гиперболического параболоида есть два семейства прямолинейных образующих: через каждую его точку проходит по образующей из одного и из другого семейства.*

Доказательство. Рассмотрим гиперболический параболоид с уравнением

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2z, \quad \text{т. е.} \quad \left(\frac{x}{a} - \frac{y}{b}\right)\left(\frac{x}{a} + \frac{y}{b}\right) = 2z. \tag{1}$$

Рассмотрим также прямую L, задаваемую системой уравнений

$$\frac{x}{a} - \frac{y}{b} = p, \quad p\left(\frac{x}{a} + \frac{y}{b}\right) = 2z, \tag{2}$$

где p — произвольное наперед заданное число.

Подставляя во второе уравнение выражение для p из первого, получаем уравнение параболоида (1). Значит, любые x, y, удовлетворяющие уравнениям (2), удовлетворяют и (1), т. е. каждая точка любой прямой L принадлежит параболоиду. Другими словами, все прямые L в нем содержатся (рис. 99).

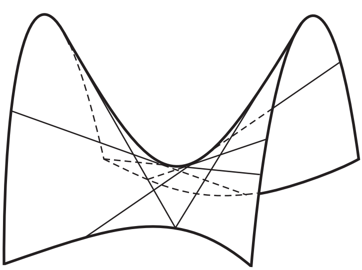


Рис. 99

Убедимся, что через каждую точку параболоида проходит прямая L. Возьмем точку параболоида $M(x_0,\ y_0,\ z_0)$ и положим в (2)

$$p = \frac{x_0}{a} - \frac{y_0}{b}.$$

Тогда при $x = x_0$, $y = y_0$, $z = z_0$ оба уравнения (2) удовлетворяются, потому что точка M на параболоиде и, значит, $2z_0 = x_0^2/a^2 - y_0^2/b^2$. Стало быть, через точку M проходит прямая L.

Следовательно, параболоид покрыт прямыми L. Они — его образующие.

Рассмотрим теперь прямые, задаваемые системами уравнений вида

$$\frac{x}{a} + \frac{y}{b} = q, \quad q\left(\frac{x}{a} - \frac{y}{b}\right) = 2z. \tag{3}$$

Для них совершенно аналогично получается тот же вывод: они покрывают параболоид — являются его образующими. □

Отметим, что каждая образующая одного семейства пересекает каждую образующую другого семейства.

В этом убеждаемся, решая совместно уравнения (2) и (3) при любых данных p и q; координаты точки пересечения:

$$x = \frac{a}{2}(q+p), \quad y = \frac{b}{2}(q-p), \quad z = \frac{1}{2}pq.$$

Прямолинейные образующие однополостного гиперболоида:

Теорема 2. *У однополостного гиперболоида есть два семейства прямолинейных образующих: через каждую его точку проходит по образующей из одного и из другого семейства.*

Доказательство. Рассмотрим однополостный гиперболоид, записав его уравнение в виде $x^2/a^2 - z^2/c^2 = 1 - y^2/b^2$, т. е.

$$\left(\frac{x}{a} - \frac{z}{c}\right)\left(\frac{x}{a} + \frac{z}{c}\right) = \left(1 - \frac{y}{b}\right)\left(1 + \frac{y}{b}\right). \tag{1}$$

Рассмотрим также систему уравнений

$$p\left(\frac{x}{a} - \frac{z}{c}\right) = q\left(1 - \frac{y}{b}\right), \quad q\left(\frac{x}{a} + \frac{z}{c}\right) = p\left(1 + \frac{y}{b}\right), \tag{2}$$

где p, q — произвольные числа, не равные одновременно нулю. При данных p, q эти уравнения определяют некоторую прямую L.

Перемножая левые части и правые части уравнений (2) и сокращая pq (если $pq \neq 0$), получаем из (2) уравнение (1). Это значит, что каждая точка прямой L, а тем самым и вся прямая, содержится в гиперболоиде (рис. 100).

Этот вывод сделан при $pq \neq 0$. Но если $pq = 0$, то по условию либо $p = 0$, а $q \neq 0$, либо наоборот. Допустим, $q = 0$, а $p \neq 0$. Тогда уравнения (2) дают систему:

$$\frac{x}{a} - \frac{z}{c} = 0, \quad 1 + \frac{y}{b} = 0$$

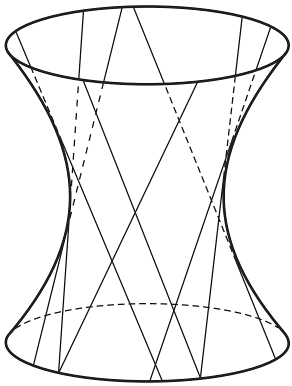


Рис. 100

и уравнение (1) выполнено. Аналогичное получим при $p = 0$, $q \neq 0$. Следовательно, и при p или q, равных нулю, прямая L содержится в гиперболоиде.

Покажем, что через каждую точку гиперболоида проходит прямая типа L. Возьмем на гиперболоиде точку $M(x_0,\ y_0,\ z_0)$. Нужно найти такие p и q, чтобы выполнялись равенства

$$\begin{aligned} p\left(\frac{x_0}{a} - \frac{z_0}{c}\right) &= q\left(1 - \frac{y_0}{b}\right), \\ q\left(\frac{x_0}{a} + \frac{z_0}{c}\right) &= p\left(1 + \frac{y_0}{b}\right). \end{aligned} \tag{3}$$

Если в первом равенстве хотя бы один из множителей при p и q отличен от нуля, то полагаем

$$p = 1 - \frac{y_0}{b}, \quad q = \frac{x_0}{a} - \frac{z_0}{c}, \tag{4}$$

и первое равенство (3) выполняется. Второе тоже выполняется, так как при подстановке в него p и q из (4) получаем уравнение (1), которое выполняется, поскольку точка $M(x_0,\ y_0,\ z_0)$ лежит на гиперболоиде.

Если первое равенство сводится к $0 = 0$, то берем второе равенство и полагаем

$$p = \frac{x_0}{a} + \frac{z_0}{c}, \quad q = 1 + \frac{y_0}{b}.$$

Здесь $q \neq 0$, так как $y_0 = b$ (поскольку в первом равенстве (4) имеем $1 - y_0/b = 0$). Если эти значения p и q подставить в первое равенство (3), то получим уравнение (1), которое выполняется, так как точка $M(x_0,\ y_0,\ z_0)$ лежит на гиперболоиде.

Итак, через каждую точку гиперболоида проходит прямая L. Он покрыт этими прямыми; они — его образующие. Совершенно такой же вывод получается для прямых, заданных системами вида

$$p\left(\frac{x}{a} + \frac{z}{c}\right) = q\left(1 - \frac{y}{b}\right), \quad q\left(\frac{x}{a} - \frac{z}{c}\right) = p\left(1 + \frac{y}{b}\right). \quad \square \tag{5}$$

Любые две образующие однополостного гиперболоида из разных семейств пересекаются, кроме того случая, когда они параллельны. Это, естественно, доказывается тем, что уравнения (2) и (5) с данными p и q решаются совместно. Тогда получается, что они разрешимы всегда, кроме того случая, когда $pq = -1$. А в этом случае прямые параллельны, как можно убедиться (ввиду неразрешимости уравнений они не пересекаются, но они могли бы скрещиваться).

Образующие гиперболоида вращения. Рассмотрим однополостный гиперболоид вращения. Он симметричен относительно плоскости своего горлового сечения и относительно всякой плоскости, проходящей через ось вращения (при каноническом уравнении эта ось — ось z; плоскость горлового сечения — плоскость $z = 0$).

Пусть O — центр гиперболоида, A — точка на горловой окружности и отрезок OA — это радиус горловой окружности, он перпендикулярен оси. Образующие гиперболоида, проходящие через точку A, перпендикулярны OA и образуют с направлением оси равные углы. Это ясно из симметрии. При отражении в плоскости, проходящей через ось и точку A, они должны переходить одна в другую так же, как при отражении в плоскости горлового сечения. Проходящая через них плоскость перпендикулярна OA; она пересекает гиперболоид по этим двум образующим.

Так как гиперболоид при вращении вокруг оси переходит сам в себя, то его образующие переходят в образующие. Одна из рассмотренных образующих при вращении дает образующие одного семейства, другая — другого семейства (рис. 101).

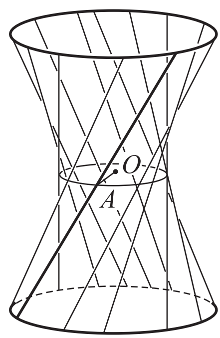


Рис. 101

Русский инженер В. Г. Шухов (1853–1939) предложил в 1896 г. конструкции из металлических балок, располагаемых так, как расположены образующие гиперболоида вращения. Такие конструкции оказались легкими и прочными; они применяются для устройства водонапорных башен и высоких радиомачт.

Можно изготовлять зубчатые колеса в форме гиперболоидов вращения с зубцами, идущими по образующим из одного семейства. Передача происходит с одного такого колеса на другое с осями, расположенными под углом. Важно, что зубцы прямые и могут быть сравнительно длинными, а усилие распределяется на всю их длину.

Поверхность с прямолинейными образующими. Поверхностей с одним семейством прямолинейных образующих — необозримое количество, но двумя семействами, оказывается, обладают только однополостные гиперболоиды и гиперболические параболоиды (не считая, конечно, плоскости — у нее не только два, а бесконечно много семейств прямолинейных образующих).

Теорема 3. *Поверхность с двумя семействами прямолинейных образующих — это, не считая плоскости, либо однополостный гиперболоид, либо гиперболический параболоид.*

Эта теорема является, можно сказать, следствием другой не менее замечательной теоремы.

Теорема 4. *Прямые, пересекающие три попарно скрещивающиеся прямые, образуют однополостный гиперболоид или гиперболический параболоид. При этом в случае гиперболоида нужно к указанным прямым присоединить еще три «особых» прямых, каждая из которых пересекает только две из данных и параллельна третьей данной прямой.*

Параболоид получается, если три данные прямые параллельны одной плоскости; если такой плоскости нет — получается гиперболоид.

Три данные прямые будут принадлежать одному семейству прямолинейных образующих получающейся поверхности, а те прямые, которые их пересекают (включая три «особых») образуют другое их семейство.

Задачи

1. Проведите во всех деталях выводы, кратко указанные в тексте:

а) о втором семействе образующих параболоида;

б) о втором семействе образующих гиперболоида;

в) о пересечении образующих параболоида;

г) о пересечении образующих гиперболоида (явно выпишите формулы для координат точки пересечения через параметры p, q; p', q' образующих одного и другого семейств и выясните, при каких параметрах образующие параллельны).

2. Выведите уравнение гиперболоида, получающегося при вращении прямой c вокруг другой прямой a; прямые скрещиваются, основание O их общего перпендикуляра на прямой a неподвижно.

3. Докажите: если три прямые попарно скрещиваются, то через каждую точку любой из них проходит прямая (и притом только одна), которая пересекает две другие, или пересекает одну и параллельна другой. А что будет, если две из трех прямых не скрещиваются, а также в других случаях взаимного расположения трех прямых? (Какие

вообще возможны случаи взаимного расположения трех прямых?)

4. Докажите: если на поверхности с двумя семействами прямолинейных образующих есть прямая, им не принадлежащая, то поверхность — плоскость.

5. Покажите, что у параболоида образующие одного семейства параллельны одной плоскости, образующие другого — параллельны другой плоскости.

6. Покажите, что образующие однополостного гиперболоида параллельны образующим его асимптотического конуса.

Часть 2
ЭЛЕМЕНТАРНАЯ ГЕОМЕТРИЯ

Во второй части мы начинаем подробное аксиоматическое построение элементарной геометрии. Оно является как бы тем стержнем, вокруг которого строится весь курс. Это построение продолжается и в третьей, и в шестой (последней) части.

Аксиоматика евклидовой геометрии, которую мы предлагаем в этой части, заведомо избыточна: она включает в себя аксиомы измерения величин — длин отрезков, мер углов, площадей простых плоских фигур, объемов простых тел. Поэтому здесь не развивается подробно теория измерения геометрических величин, а ее изложение откладывается до шестой части. Там же будет дана другая, более экономная аксиоматика евклидовой геометрии.

Здесь же, в первых двух главах, исходя из предложенной аксиоматики, выведены основные теоремы элементарной геометрии (или указан путь, на котором они могут быть легко получены).

Третья глава посвящена более специальным вопросам элементарной геометрии.

Следует отметить, что в этой части подробно обсуждаются различные возможные подходы к определениям основных понятий школьной геометрии — отрезка, прямой, угла, многоугольника, многогранника, геометрического тела и т. д.

Глава I
АКСИОМЫ ГЕОМЕТРИИ

§ 1. Общее понятие об основаниях геометрии

Основания геометрии образуются теми ее понятиями и положениями, исходя из которых можно ее развивать путем чисто логических рассуждений.

Положение, принимаемое без доказательства в качестве исходного, отправного для данной теории, в частности для геометрии, называется *аксиомой.* Аксиомы говорят об основных понятиях теории.

Совокупность аксиом, лежащих в основаниях теории, называют *аксиоматикой* этой теории; говорят также: «система аксиом».

К аксиоматике можно относить и основные понятия, и тогда слова основания геометрии и аксиоматика геометрии означают одно и то же. К основаниям можно еще относить доказательства исходных теорем и обоснование важнейших понятий, таких как, скажем, понятие площади.

Доказательством называется, вообще, убедительное рассуждение. Положение (утверждение) теории, которое доказывается или подлежит доказательству, называется *теоремой.* Обычно, тем более в школьном изложении геометрии, в доказательствах используют наглядные соображения. Но при строго логическом изложении это нужно исключить, и доказательство должно опираться только на уже установленные положения. А так как они, в свою очередь, тоже должны на чем-то основываться, то приходим к тому, что в основе должны лежать некоторые положения, принимаемые без доказательства, т. е. аксиомы.

Аналогичное выполняется для понятий и их определений. Обычно определение состоит в том, что определяемое понятие разъясняется через другие, можно сказать, к ним сводится. Но нельзя сводить одни понятия к другим до бесконечности. Поэтому должны быть исходные понятия, которые принимаются без предварительных определений. Все, что от них требуется, высказывается в аксиомах (поэтому говорят, что аксиомы служат *«скрытыми» определениями* основных понятий).

Понятия могут быть двух типов: одни описывают объекты, такие как «точка», «прямая», «круг» и т. п., другие описывают отношения объектов друг к другу, такие как «прямые пересекаются», «точка лежит внутри круга» и т. п.

Основные понятия при изложении геометрии можно выбирать поразному, например, мы примем за основное понятие «отрезок», тогда как чаще принимают «прямую». В качестве аксиом тоже можно брать разные положения геометрии; такие возможности мы рассмотрим специально в части 6.

Аксиоматика излагается так, что сначала перечисляются основные понятия, а потом формулируются аксиомы. Например, примем следующие основные понятия.

О б ъ е к т ы: *точки, прямые.*

О т н о ш е н и е: точка *принадлежит* прямой,

Сформулируем **аксиому:** *каждые две точки принадлежат некоторой прямой* (или, как чаще говорят, через каждые две точки проходит прямая).

Геометрия возникла из практики, в частности из измерения земли; само слово «геометрия» означает «землемерие». Греческий ученый Евдем Родосский (4 в. до н. э.) писал: «Геометрия была открыта египтянами и возникла из измерения земли. Это измерение было им необходимо вследствие разливов Нила, постоянно смывавших границы (участков земли). Нет ничего удивительного в том, что эта наука, как и другие, возникла из потребностей человека. Всякое возникающее знание из несовершенного состояния переходит в совершенное. Зарождаясь путем чувственного восприятия, оно постепенно становится предметом нашего рассмотрения и, наконец, делается достоянием разума».

Так геометрические знания, зародившиеся из чувственного восприятия в практике, стали достижением разума в логической системе «Начал» Евклида (ок. 300 г. до н. э.). Однако с современной точки зрения эта система далеко не совершенна. Принятое теперь понимание оснований геометрии сложилось лишь к концу XIX столетия.

Согласно происхождению геометрии, ее основания лежат фактически в практике: для планиметрии, можно сказать, — в измерениях земли. Соответственно, наиболее естественные логические основания геометрии, содержащиеся в ее аксиоматике, должны возможно ближе выражать ту же практику, лишь представляя ее в идеализированном виде. Так мы и подойдем к основаниям планиметрии. К основаниям стереометрии мы обратимся, опираясь на планиметрию.

Замечания. 1. Слово «аксиома» происходит от греческого и означает «достойное признания»; прежде и понимали «аксиому» как положение, достойное признания ввиду его очевидности, не требующее доказательства, безусловное. В этом смысле и теперь говорят, например, о моральных аксиомах.

Аксиомы геометрии тоже толковали как «не требующие доказательства по очевидности». Например, в известном учебнике геометрии Киселева так и было написано: «Аксиомы. Так называют истины, которые вследствие своей очевидности принимаются без доказательства»[1]. Но понимание это изменилось; в «Словаре русского языка» дается определение: «Аксиома — положение, принимаемое без доказа-

[1] *Киселев А. П.* Элементарная геометрия. М., 1914. С. 1.

тельства в качестве исходного, отправного для данной теории. Неоспоримая истина, совершенно очевидное утверждение».

Таким образом, прежнее значение слова «аксиома» толкуется теперь как вторичное. В науке же «аксиома» понимается всегда в смысле первого определения; его мы и воспроизвели в начале нашего изложения.

Соответственно, от аксиом в принципе не требуется ничего, кроме того, чтобы они давали основание теории. Но в геометрии, поскольку она неразрывно связана с наглядным содержанием, пренебрегать очевидностью аксиом не следует.

2. Термином «геометрия» обозначают теперь многие теории, порой довольно далекие от первоначальной элементарной геометрии. Но мы, говоря о геометрии, будем иметь в виду обычные — евклидовы — планиметрию и стереометрию, пока не перейдем, указав на это явно, к другим теориям, включаемым в геометрию (см. части 3 и 6).

§ 2. Основные понятия аксиоматики планиметрии

Согласно общему плану, указанному в § 1, изложение аксиоматики начинается с перечисления основных понятий — объектов и отношений. Мы принимаем за

Основные объекты: 1) *точки*, 2) *отрезки*.

Основные отношения: 1) точка *служит* концом отрезка, 2) точка *лежит* на отрезке, 3) два отрезка *равны* друг другу, или, как мы будем также говорить, один отрезок *равен* другому.

Точки обозначаются, как обычно: A, B, M и т. п., отрезки — a, b и т. п. Если точка C лежит на отрезке a, то кратко пишем: C *на* a (см. рис. 1)[2].

Рис. 1. Точки A, B — концы отрезка a, C — лежит на a

Точки, лежащие на отрезке, а также его концы считаются точками этого отрезка, т. е. принимаем определение: точка A *принадлежит* отрезку a, если она лежит на нем или служит его концом. Это обозначаем так: $A \in a$ (не смешивать с «A на a»!).

Теперь, пользуясь основными понятиями, определим некоторые

---

[2] Говорят также: «точка лежит внутри отрезка»; но всегда говорят: «на прямой», «на стороне треугольника» и т. п. Выражение «на отрезке» короче, чем «внутри отрезка», и более соответствует наглядному представлению, когда отрезок рассматривается сам по себе, а не как часть прямой.

другие понятия, относящиеся к отрезкам (не придерживаясь принятой в школе формы определения с глаголом «называется»).

1. Отрезок a *содержится* в отрезке b (в записи $a \subset b$), если все его точки являются также точками отрезка b (рис. 2, *а*).

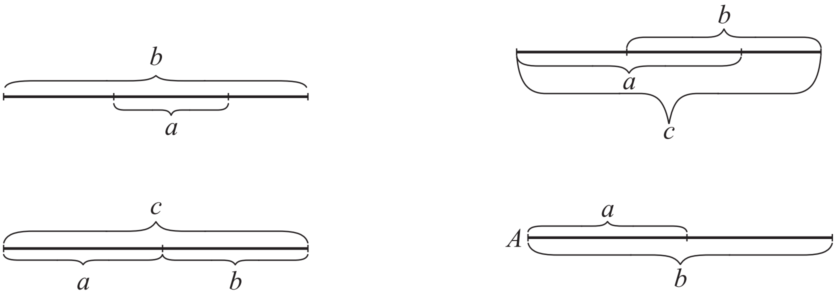


Рис. 2. *а* — a содержится в b; *б* — a и b образуют c;
в — a и *б* составляют c; *г* — a и b налегают друг на друга

2. Отрезки a, b *образуют отрезок* c (в записи $c = a \cup b$), если они содержатся в c и у c нет точек, не принадлежащих им (рис. 2, *б*). Дословно так же определяется, что несколько отрезков образуют один отрезок.

3. Отрезок c *составлен* из отрезков a, b, если $c = a \cup b$ и отрезки a, b не имеют общих точек, кроме одного общего конца (рис. 2, *в*). Вообще, отрезок c составлен из нескольких данных отрезков, если они образуют его и попарно не имеют общих точек, кроме концов.

4. Отрезок a *отложен вдоль* отрезка b от его конца A, если у этих отрезков есть общий конец A и один из них содержится в другом. Мы говорим также, что такие отрезки *налегают* друг на друга (рис. 2, *г*).

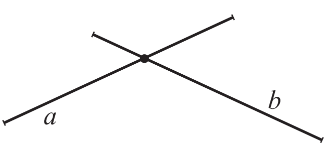


Рис. 3. a и b пересекаются

5. Два отрезка *пересекаются*, если у них есть единственная общая точка, лежащая на них обоих (рис. 3).

Подчеркнем, что при аксиоматическом изложении основные понятия заранее не определяются; все, что от них требуется, должно быть высказано в аксиомах. Но смысл основных понятий и содержание аксиом отражают то, откуда они возникли. Наши основные объекты можно определить из их исходного наглядного смысла.

Точка — это мысленный образ предельно точно указанного места,

так что в нем уже не различаются разные места. По Евклиду «точка — это то, что не имеет частей».

Отрезок — это мысленный образ туго натянутой нити.

Однако такие определения носят «предматематический» характер, и их не следует смешивать с математическими определениями, поскольку в них участвуют не математические понятия, такие как «место» или «нить».

Аксиомы планиметрии мы делим на две группы:

а) *Линейные аксиомы.* В них не присутствует представление о плоскости, так что они могли бы относиться к точкам и отрезкам, лежащим на одной прямой. Поэтому мы и называем их *линейными.*

б) *Плоскостные аксиомы.* Они касаются фигур, не укладывающихся на прямой.

Говоря здесь о прямой и плоскости, мы понимаем их в наглядном смысле; понятие о них в нашу аксиоматику не включается, так что сказанное представляет собой только пояснение к разделению аксиом на «линейные» и «плоскостные».

Поскольку аксиомы должны давать исчерпывающее основание для вывода теорем без всяких ссылок на наглядные представления, то в них указываются, в частности, и такие свойства основных объектов, которые при наглядном понимании представляются настолько очевидными, что их не стоит упоминать, как, например, первая из линейных аксиом, с которой начинается следующий параграф.

Линейные аксиомы мы делим, в свою очередь, на две подгруппы:

а) «Аксиомы связи» — связи точек и отрезков. Эти аксиомы касаются только отношений точек и отрезков: точка служит концом отрезка и точка лежит на отрезке.

б) Аксиомы равенства и измерения отрезков.

Замечания. 1. Мы излагаем основания планиметрии — геометрии на плоскости. Само понятие плоскости в них не упоминается; в них плоскость — это, так сказать, та среда, то пространство, где выполнены высказываемые далее аксиомы и соответственно «разыгрывается», имеет место основанная на них геометрия.

2. В обыденной речи говорят не «отрезок», а «прямая» — «проведем прямую», «иди по прямой» и т. п. Но при этом никто не имеет в виду «бесконечную прямую во всей ее бесконечности», — в отличие от того, как это принято теперь в геометрии. Понятие «конечной прямой» — отрезка — первично и берется из практики, а понятие «бесконечной прямой» возникает из *представления* о возможности продолжать отрезок за оба конца без ограничения. В современном изложении начал

геометрии все переворачивается: бесконечная прямая принимается за нечто первичное, а отрезок определяют как часть прямой.

Однако в геометрии прямые в их актуальной бесконечности почти не встречаются, а всюду фигурируют отрезки. Даже параллельные прямые, появляясь в определении и в аксиоме о параллельных, дальше не нужны: всегда имеют дело не с ними, а с параллельными отрезками. Далее. Отрезок изображается, например, чертой на бумаге, но подобное изображение (модель) бесконечной прямой уже невозможно: бесконечная — значит, уходящая за всякие пределы Вселенной. В учебниках и преподавании геометрии говорят о прямой, проведенной на чертеже, но на нем проводят конечную прямую — отрезок— с той точностью, с какой позволяют средства черчения. Поэтому, строго говоря, в преподавании допускается путаница и некоторый обман.

Наконец, можно заметить, что в «Началах» Евклида прямая понимается как конечная. Еще Лобачевский имел в виду конечную прямую, когда писал: «Величина прямой линии определяется сравнением ее с другой»[3]. В учебнике Киселева различаются «конечная» и «бесконечная» прямые[4]. Параллельные — это конечные прямые, которые не пересекаются, как их ни продолжать.

§ 3. Линейные аксиомы связи и их первые следствия

Как уже сказано, эти аксиомы касаются только отношений точек и отрезков.

I$_1$ (аксиома существования). *Существует хотя бы один отрезок; у каждого отрезка есть два и только два конца; кроме того, на каждом отрезке лежит хотя бы одна точка.*

I$_2$ (аксиома проведения отрезка). *Каждые две точки служат концами отрезка, и притом только одного. Другими словами: каждые две точки можно соединить отрезком, и притом единственным.*

Отрезок с концами A, B обозначаем обычно AB (или BA).

I$_3$ (аксиома деления отрезка). *Всякая точка, лежащая на отрезке, делит его на два отрезка, т. е. если C на AB, то AB составлен из AC и BC.*

I$_4$ (аксиома соединения отрезков). *Два отрезка, имеющие две общие точки, образуют один отрезок; его концами служат два из их концов.*

[3] См.: Об основаниях геометрии. М.; Л., 1956. С. 32.

[4] *Киселев А. П.* Элементарная геометрия. М.: Просвещение, 1980. С. 4.

То есть либо концы этого отрезка — это концы одного из данных отрезков (рис. 4, а), либо один из его концов — это конец одного из двух данных отрезков, а другой — конец другого отрезка (рис. 4, *б*). Конечно, может быть, что одна пара концов или даже обе совпадают (рис. 4, *в*).

Эта аксиома выражает тот простой факт, что если провести через две данные точки несколько разных отрезков, то они вместе образуют один отрезок.

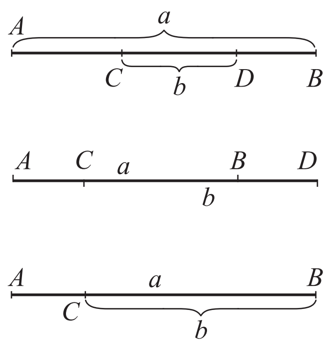


Рис. 4.
а — AB и CD образуют AB;
б — AB и CD образуют AD;
в — AB и CB образуют AB

Первые теоремы о связи точек и отрезков. Эти теоремы устанавливают совершенно ясные наглядно свойства отрезков, но тем поучительнее их скрупулезное доказательство, основанное только на аксиомах.

Теорема 1. *Точка, лежащая на отрезке, не может быть его концом.* (Стало быть, и конец не может лежать на отрезке.)

Доказательство. Если точка C на AB, то по аксиоме I_3 она делит отрезок AB на AC и BC. Это, в частности, означает, что точка C отлична от A и от B. Но если бы C была концом отрезка AB, то она совпадала бы либо с A, либо с B (так как у отрезка только два конца). Противоречие. □

Теорема 2. *Если* $C \in AB$ *и* $D \in AB$, *то* $CD \subset AB$. *Если при этом* M *на* CD, *то и* M *на* AB.

Доказательство. Пусть C и $D \in AB$ и M на CD. Если точки C и D совпадают с A и B, то по аксиоме $\mathbf{I}_2$ отрезок CD совпадает с AB, так что M на AB.

Допустим, что только одна из точек C, D совпадает с A или B: скажем, D совпадает с A, а C отлична от B. Тогда, во-первых, отрезок CD — тот же, что AC, так что M на AC. Во-вторых, так как точка C принадлежит AB, но отлична от A и B, то C на AB. Следовательно (по аксиоме деления $\mathbf{I}_3$), точка C делит AB на AC и BC, и так как M на AC, то 1) $M \in AB$; 2) M отлична от A и C (по теореме 1); 3) M отлична от B, так как у отрезков AC, BC общая только точка C (по аксиоме деления). Следовательно, M на AB, что и требовалось доказать.

Пусть теперь обе точки C, D отличны от концов отрезка AB и,

значит, лежат на AB. Точка C делит AB (по аксиоме деления), и D лежит либо на AC, либо на BC. Пусть, например, D на AC. Тогда если M на CD, то, по предыдущему выводу, M на AC, а потому также на AB (опять по предыдущему выводу), что и требовалось доказать. □

Из теоремы 2 следует

Теорема 3. *Отрезок определяется своими точками: если у отрезков a и b одни и те же точки, т. е. $a \subset b$ и $b \subset a$, то это один и тот же отрезок.*

(У отрезка два рода точек: те, которые лежат на нем, и концы. Поэтому надо доказать, что у отрезков a и b концы одни и те же. Ведь логически мыслимо, что конец одного отрезка оказывается внутри другого; и надо вывести из аксиом, что это невозможно, когда $a \subset b$ и $b \subset a$.)

Доказательство. Пусть $a \subset b$ и $b \subset a$. То, что $a \subset b$, означает, в частности, что концы отрезка a содержатся в b, и потому, в силу теоремы 2, точки, лежащие на a, лежат также на b. Но так как $b \subset a$, то, точно так же, точки, лежащие на b, лежат на a. Таким образом, точки, лежащие на отрезках a и b, одни и те же. Значит (в силу теоремы 1), их концы одни и те же; так что, по аксиоме $\mathbf{I}_2$, a и b — это один отрезок. □

Теорема 4. *Если отрезки с общим концом имеют еще одну общую точку, то они налегают друг на друга.*

Доказательство. Пусть отрезки AB, AC имеют общую точку помимо A. Тогда, по аксиоме $\mathbf{I}_4$, они образуют один отрезок, концами которого служат две из точек A, B, C, т. е. этот отрезок либо AB, либо AC, либо BC. Но в последнем случае выходило бы, что A на BC, и, по аксиоме деления отрезка, AB и AC не имели бы общих точек, кроме A. Поэтому отрезки AB, AC образуют вместе либо AB, либо AC. В первом случае $AB \supset AC$, во втором — $AC \supset AB$, т. е. отрезки налегают один на другой, что и требовалось доказать. □

§ 4. Аксиомы равенства и измерения отрезков

Здесь мы формулируем те аксиомы, в которых участвует отношение равенства отрезков и вводится длина отрезка. Длину отрезка измеряют, сравнивая отрезок с тем отрезком, который принят за масштаб, и находят численное значение длины в данном масштабе, т. е. при данной единице измерения. Теоретически измерение предполагается абсолютно точным, так что каждому отрезку соответствует определенная длина, выражаемая вещественным (действительным) числом. Это мы

принимаем здесь как аксиому. А те аксиомы, которые обеспечивают абсолютно точное измерение отрезков, и сам идеальный процесс измерения будут рассмотрены позже в части 6 «Основания геометрии».

Сформулируем аксиомы:

II$_1$ (аксиома откладывания отрезка). *Вдоль любого отрезка от любого из его концов можно отложить отрезок, равный любому данному, и притом только один.*

II$_2$ (аксиома длины). *Если некоторому отрезку e отнесено число единица, то каждому отрезку можно сопоставить положительное число — его численную длину в масштабе e — так, чтобы были выполнены условия:*

а) *равные отрезки имеют одну и ту же длину;*

б) *если C на AB, то длина отрезка AB равна сумме длин отрезков AC, BC.* (Здесь, как и дальше, мы говорим «длина» вместо «численная длина в данном масштабе».)

II$_3$ (аксиома существования отрезка данной длины). *При любом данном масштабе e при всяком положительном числе l существует отрезок с длиной, равной l.*

Обозначение. В этой главе (и только в ней!) равенство отрезков будем обозначать знаком $\cong$. Длину отрезка (в каком-либо масштабе) будем обозначать как модуль: $|AB|$ и т. п.

В этом обозначении условие 2 аксиомы **II**$_2$ запишется так:

Если C на AB, то $|AB| = |AC| + |BC|$.

Напомним, что длина отрезка называется также *расстоянием* между его концами. Расстояние от точки до нее самой полагается равным нулю, т. е. $|AA| = 0$.

Замечание. Сама по себе длина есть ***величина*** и не зависит от масштаба; она получает численное значение при измерении тем или иным масштабом. Например, 3 см и 30 мм — это одна и та же длина, а численные значения разные. Об этом не надо забывать, когда мы говорим о «длине», подразумевая, однако, ее численное значение при каком-либо масштабе. (О понятии величины см. часть 6.)

Аксиому II$_2$ можно пересказать, введя сначала определение.

Численной длиной в масштабе e называется функция отрезка, положительная, равная единице для отрезка e и обладающая свойствами а), б) из аксиомы II$_2$, т. е.

а) она имеет одно и то же значение для равных отрезков;

б) ее значение для любого отрезка равно сумме ее значений для отрезков, из которых он составлен.

Аксиома II_2 утверждает: При данном отрезке e существует функция с указанными свойствами. То есть каждый отрезок имеет численную длину в данном масштабе e.

Эта аксиома о существовании численной длины дополняется теоремой о ее единственности.

Теорема 1. *При данном масштабе e указанная функция единственна, т. е. каждый отрезок имеет единственную вполне определенную численную длину.*

Кроме того, выполняется следующая теорема о замене масштаба, из которой вытекает теорема 1.

Теорема 2. *При перемене масштаба численная длина изменяется на положительный множитель, именно, если e, e' — два масштаба и $l(a)$, $l'(a)$ — длины любого отрезка a в этих масштабах, то*

$$l'(a) = c \cdot l(a), \quad c = \mathrm{const} > 0. \tag{1}$$

Если взять $a = e$, то так как $l(e) = 1$, получаем, что $c = l'(e)$, и если взять $a = e'$, то получаем $c = 1/l(e')$. Таким образом, из (1)

$$l'(a) = l'(e)l(a) = \frac{l(a)}{l(e')}.$$

То есть при переходе к «новому» масштабу e' от «старого» e длина делится на длину отрезка e' в «старых» единицах e.

Теорема 1 является частным случаем теоремы 2, соответствующим тому, что масштаб e «меняется» на тот же e. Тогда как $l(e) = 1$, так и $l'(e) = 1$. Поэтому, полагая в (1) $a = e$, получим $c = 1$ и $l'(a) = l(a)$, т. е. длина любого отрезка та же самая, как и утверждает теорема 1. Теорема 1 доказана. $\square$

Все это можно понять так, что длина может измеряться не только разными масштабами, но и разными способами. Тогда, как утверждает теорема 2, если выполнены условия а), б), длина зависит только от масштаба.

Теорему 2 мы докажем дальше — в конце этого параграфа, и читатель, которому ее доказательство представится трудным, может принять ее либо как аксиому II_4 о замене масштаба, либо как теорему, доказательство которой пока откладывается.

А теперь докажем некоторые теоремы о длине отрезков.

По аксиоме II_2 у равных отрезков длины равны; верно также обратное:

Теорема 3. *Если длины двух отрезков (в каком-либо данном масштабе) равны, то и сами эти отрезки равны.*

Доказательство. Пусть у отрезков AB, CD длины равны: $|AB| = |CD|$. Отложим вдоль AB отрезок AM, равный CD; по аксиоме II_1 это возможно. По аксиоме II_2, $|AM| = |CD|$, так что $|AB| = |AM|$.

Если AM совпадают с AB, то это значит, что отрезок CD равен AB, как утверждает эта теорема.

Допустим, что AM не совпадает с AB. Но так как AM налегает на AB, то либо AM содержится в AB, так что M на AB, либо, напротив, $AB \subset AM$ и B на AM. В первом случае, по аксиоме II_2, $|AB| = |AM|+|MB|$. Тогда так как $|MB| > 0$ (по аксиоме II_2), то $|AB| > |AM|$, вопреки тому, что $|AM| = |AB|$. Стало быть, M не может лежать на AB.

Если же B на AM, то $|AM| = |AB| + |BM|$, так что $|AM| > |AB|$, вопреки тому, что $|AM| = |AB|$. Стало быть, B не может лежать на AM.

Остается лишь та возможность, что отрезки AM и AB совпадают, и так как $AM \cong CD$, то $AB \cong CD$, что и требовалось доказать. $\Box$

Из теоремы 3 вместе с аксиомой II_2 следует

Теорема 3а. *Равенство отрезков равносильно равенству их численных длин (в любом данном масштабе).* $\Box$

Равенство чисел рефлексивно, симметрично и транзитивно. Поэтому из теоремы 3а вытекает

Следствие. *Отношение равенства отрезков рефлексивно, симметрично и транзитивно*[5].

Впрочем, равенство отрезков симметрично по понятию, ведь оно выражено так: два отрезка равны *друг другу.* Ввиду этого транзитивность означает, что два отрезка, равные одному и тому же отрезку, равны друг другу. Рефлексивность означает, что всякий отрезок равен самому себе.

Теорема 4. *Между точками M любого данного отрезка AB и числами $x \in [0,\ |AB|]$, заключенными между нулем и длиной $|AB|$ отрезка AB, имеется взаимно однозначное соответствие, при котором каждому x отвечает такая точка M, что $|AM| = x$. При этом если $0 < x_1 < x_2$, то для соответствующих точек M_1, M_2 будет $AM_1 \subset AM_2$ и $|M_1M_2| = x_2 - x_1$.*

---

[5] Пусть R — некоторое отношение каких-либо объектов a, b, c и т. д., так что пишем aRb и т. п. Рефлексивность означает, что aRa для всякого a, симметричность — что если aRb, то bRa; транзитивность — что если aRb и bRc, то aRc.

Доказательство. Каждому отрезку AM отвечает определенная длина $|AM|$ (по аксиоме II_2). Если M на AB, то, по аксиоме II_2, $|AB| = |AM| + |MB|$, так что $|AM| < |AB|$.

Вместе с тем по аксиоме II_3 для каждого $x > 0$ существует отрезок длины x. По аксиоме II_1 равный ему отрезок AM можно отложить вдоль AB, и $|AM| = x$ по аксиоме II_2. Такой отрезок AM только один. Если $x < |AB|$, то $AM \subset AB$ и точка M лежит на AB. (Если бы было $AB \subset AM$ или $AB \cong AM$, то из аксиомы II_2 получалось бы, что $|AB| \leqslant |AM|$, вопреки тому, что $|AM| = x < |AB|$.)

Итак, каждой точке M на AB соответствует длина $x = |AM| < |AB|$, и обратно, если $0 < x < |AB|$, то на AB есть единственная точка M такая, что $|AM| = x$. Таким образом, между точками M и числами x из промежутка $(0,\ |AM|)$ имеется взаимно однозначное соответствие.

При этом если M_1 на AM_2, то, по аксиоме II_2,, $|AM_2| = |AM_1| + |M_1M_2|$. Тогда $|AM_1| < |AM_2|$ и $|M_1M_2| = |AM_2| - |AM_1| = x_2 - x_1$.

Покажем, что и обратно: если $x_1 < x_2$ и точки M_1, M_2 таковы, что $|AM_1| = x_1$, $|AM_2| = x_2$, то M_1 на AM_2, и, по доказанному, $|M_1M_2| = |AM_2| - |AM_1| = x_2 - x_1$.

Действительно, допустим вопреки утверждению, что M_1 не на AM_2. Тогда, поскольку M_2 делит AB на AM_2 и BM_2, M_1 на BM_2, и, по только что доказанному, $|BM_2| > |BM_1|$. Но так как

$$|BM_2| + |M_2A| = |AB| = |BM_1| + |M_1A|,$$

то, выходит, $x_2 = |AM_2| < |AM_1| = x_1$, вопреки тому, что $x_2 > x_1$. Следовательно, M_1 на AM_2, что и требовалось доказать. $\square$

Доказанная теорема означает, что порядок расположения точек на отрезке — такой же, как порядок чисел от нуля до длины отрезка (в каком угодно данном масштабе).

Доказательство теоремы 2 о замене масштаба. Согласно теореме 3а равенство отрезков равносильно равенству их длин. Кроме того, по аксиоме II_3 при данном масштабе существуют отрезки любой данной длины, т. е. длины принимают все возможные положительные значения.

Таким образом, при данном масштабе каждому классу равных отрезков соответствует положительное число, и обратно — каждому положительному числу отвечает класс отрезков такой длины. Короче, между положительными числами и классами равных отрезков имеется взаимно однозначное соответствие. Каждый класс равных отрезков может быть представлен положительным числом. Поэтому если $l(a)$ —

длина в масштабе e, а $l'(a)$ — в масштабе e', то $l'(a)$ зависит только от $l(a)$, так что можно написать:

$$l'(a) = f(l(a)), \tag{2}$$

и $f(l(a))$ есть просто функция положительного числа $x = l(a)$.

Если отрезок c складывается из a и b, то по условию б) аддитивности аксиомы II_2

$$l(c) = l(a) + l(b) \quad \text{и} \quad l'(c) = l'(a) + l'(b).$$

То есть, пользуясь (2), получаем $f(l(a) + l(b)) = f(l(a)) + f(l(b))$. Или, полагая $l(a) = x,\ l(b) = y$, имеем

$$f(x + y) = f(x) + f(y).$$

Итак, длина отрезков в новом масштабе представляется как функция $f(x)$ с условиями:

а) она определена для всех $x > 0$,

б) она положительна: $f(x) > 0$,

в) аддитивна, т. е. $f(x + y) = f(x) + f(y)$.

Оказывается — и это мы сейчас докажем — что функция с такими условиями имеет простой вид: $f(x) = cx$, где $c = \mathrm{const} > 0$. А так как при $x = l(a)$ $f(x) = f(l(a)) = l'(a)$, то выходит, что

$$l'(a) = cl(a),$$

что и утверждает теорема 2 о замене масштаба.

Таким образом, для доказательства этой теоремы нам достаточно доказать следующую теорему.

Теорема 5. *Если функция $f(x)$, определенная для всех $x > 0$, положительна и аддитивна, т. е.*

$$f(x + y) = f(x) + f(y), \tag{1}$$

то $f(x) = cx$, где $c = \mathrm{const} > 0$.

Доказательство. Пусть $f(x)$ — функция с указанными свойствами. Из (1) следует, что $f(2x) = f(x) + f(x) = 2f(x)$. Отсюда $f(3x) = f(2x) + f(x) = 2f(x) + f(x) = 3f(x)$, и продолжая (по индукции), получим при любом натуральном n

$$f(nx) = nf(x). \tag{2}$$

Полагая здесь $x = \frac{m}{n}y$ (m, n — натуральные), получим

$$f(my) = f\left(n\frac{m}{n}y\right) = nf\left(\frac{m}{n}y\right);$$

т. е.

$$f\left(\frac{m}{n}y\right) = \frac{1}{n}f(my).$$

А так как из (2) $f(my) = mf(y)$, то

$$f\left(\frac{m}{n}y\right) = \frac{m}{n}f(y).$$

Таким образом, полагая $y = 1$, получим, что при любом рациональном $r = m/n$

$$f(r) = rf(1) = cr, \quad c = f(1). \tag{3}$$

Пусть $x_1 > x_2$, так что $x_1 = x_2 + y$, $y > 0$. Тогда

$$f(x_1) = f(x_2 + y) = f(x_2) + f(y) > f(x_1).$$

Поэтому если x — любое вещественное число и r_1, r_2 — такие рациональные числа, что

$$r_1 < x < r_2,$$

то

$$f(r_1) < f(x) < f(r_2).$$

Поэтому из (3)

$$cr_1 < f(x) < cr_2.$$

Беря последовательность чисел r_1, r_2 $(r_1 < x < r_2)$, сходящихся к x, получим отсюда (в пределе), что $f(x) = cx$, что и требовалось доказать. □

§ 5. Прямая. Понятие фигуры

Прямая. Любой отрезок AB можно неограниченно продолжать за каждый из его концов, откладывая вдоль него отрезки AM, BN все большей и большей длины (что возможно в силу аксиом II_3, II_1).

Так мы получаем прямую, представляя ее себе как неограниченно продолжаемый в обе стороны отрезок или как результат такого продолжения — как всю бесконечную прямую. Выразим это представление в виде определения:

Прямой называется объединение всех отрезков, содержащих какие-либо две данные точки.

Если это точки A, B, то соответствующая прямая обозначается (AB).

Напомним, что объединением некоторых фигур называется фигура, содержащая все точки этих фигур и никакие другие. Вместо слова «фигура» можно говорить «множество точек». Так что объединение фигур определяется так же, как объединение любых множеств.

Теорема 1. *Через каждые две точки проходит прямая, и притом только одна.*

Доказательство. То, что через каждые две точки проходит прямая. содержится в определении прямой (поскольку по аксиоме I_2 есть отрезок, соединяющий эти точки). Остается доказать, что прямая, содержащая две данные точки, только одна. Для этого достаточно убедиться, что если точки C, D принадлежат прямой (AB), то она является также прямой (CD) (поэтому если у двух прямых есть две общие точки M, N, то они являются прямыми (MN), т. е. это одна прямая).

Итак, пусть C, D — две точки прямой (AB). Это значит, что есть отрезки, содержащие точки A, B, C и A, B, D. По аксиоме I_4 они образуют один отрезок a; он содержит как точки A, B, так и C, D.

Пусть теперь b — какой-либо из отрезков, образующих прямую (AB), т. е. содержащий точки A, B. По той же аксиоме I_4 он образует с отрезком a один отрезок c, который тем самым содержит точки C, D и, стало быть, содержится в прямой (CD). Итак, каждый из отрезков, образующих прямую (AB), содержится в прямой (CD). Следовательно, $(AB) \subset (CD)$.

Но, беря отрезок b из тех, что образуют прямую (CD), мы точно так же убедимся, что $(CD) \subset (AB)$. Следовательно, (AB) и (CD) — это одна и та же прямая, что и требовалось доказать. □

Луч (полупрямая). Луч мы представляем себе как отрезок, неограниченно продолжаемый за один из его концов или как результат такого продолжения — бесконечный луч. Это приводит к определению.

Лучом называется объединение всех отрезков с общим концом, проходящих через общую точку. Общий конец этих отрезков называется *началом* луча. О точке луча, не являющейся его концом, говорят, что

она *лежит* на луче. Луч с началом O, проходящий через точку A, обозначается OA.

Если точка B лежит на луче OA, то луч OB совпадает с лучом OA.

Всякая точка прямой делит ее на два луча: служит общим началом двух лучей, не имеющих кроме нее других общих точек и образующих вместе всю прямую.

Вывод этих двух утверждений мы оставляем читателю в качестве задачи. (В первом случае можно воспользоваться тем, что отрезки OA, OB налегают друг на друга; во втором можно опираться на аксиому о делении отрезка.)

Несколько другие определения прямой и луча. Данные выше определения прямой и луча можно пересказать, пользуясь понятием геометрического места или, что то же, множества точек.

Прямая есть геометрическое место (*множество*) *точек, принадлежащих отрезкам, содержащим какие-либо две данные точки.*

Геометрическим местом точек с данным условием (подчиненных какому-либо условию) называется фигура, содержащая все точки с этим условием и не содержащая никаких других.

Прямая (AB) как фигура, служащая объединением всех отрезков, содержащих точки A, B, тем самым содержит все точки, каждая из которых принадлежит вместе с A, B одному отрезку, и не содержит никаких других точек. Стало быть, она является геометрическим местом этих точек. *Оба определения прямой равносильны.*

Совершенно также можно дать определение луча.

Луч — это множество (*геометрическое место*) *точек, принадлежащих отрезкам, имеющим общий конец и содержащим данную точку.*

Как уже сказано, «геометрическое место точек» и «множество точек» означает одно и то же, так же как и слово «фигура». В настоящее время принято говорить «множество точек», а термин «геометрическое место» встречается только в школьных учебниках. Мы будем употреблять термин «фигура» хотя бы потому, что он короче, и будем говорить о множестве точек, когда явно имеется в виду какое-либо условие, налагаемое на точки. (Вспомним определения окружности, эллипса, гиперболы и т. д.)

Замечание. Прежде, до появления теории множеств, около 100 лет назад, термин «множество точек» не употреблялся. Фигура мыслилась как нечто целое, не состоящее из точек, но как «место точек» — «место», на котором или в котором лежат точки. Теперь говорят, что

«фигура состоит из точек». Но состоят ли отрезок или окружность из точек подобно тому, как куча песка состоит из песчинок? Такое представление спорно — оно и не нужно для геометрии. Нужно только, что фигура определяется своими точками, как об этом уже говорилось в связи с заданием фигур уравнениями. Остальное — дело склонностей, кому интуитивно ближе одно или другое представление о фигурах.

Некоторое время назад в учебниках геометрии было введено понятие множества — фигура определялась как множество точек и даже говорилось, что теперь отказались от старого представления фигуры как места точек, но для этого не было точных логических оснований, а только интуитивное представление.

Аксиоматика фигуры. Геометрия, можно сказать, — это наука о фигурах. Поэтому для ее оснований нужно определить, что называется в геометрии фигурой. Простейшие фигуры — точки и отрезки — мы ввели в аксиоматике как основные объекты. Но выше мы определили прямые, лучи и могли определить окружности, эллипсы и др. Однако в эти определения входит понятие фигуры или множества точек. Если определять, как это делалось, фигуру как вообще множество точек, то следует спросить: что такое множество? Обычный ответ таков: множество — это вообще совокупность, которая состоит из точек. Однако мы уже заметили, что предположение, скажем, об отрезке как состоящем из точек спорно и не нужно для геометрии. В понятии фигуры важно лишь то, что фигура определяется своими точками.

Итак, для того чтобы иметь полные аксиометрические основания геометрии, нужно включить в них аксиомы, определяющие понятие фигуры. (Читатель, которому излагаемая далее аксиоматика фигуры покажется трудной для понимания, может ее пропустить, оставаясь с обычным интуитивным представлением о фигурах. К аксиоматике фигуры мы еще вернемся в последней части курса вместе с более глубоким изложением оснований геометрии).

В аксиоматике фигуры принимаются основные объекты: *точки* и *фигуры* — и основное отношение: точка *принадлежит* фигуре, или, что то же: фигура *содержит* точку.

Принадлежность точки A фигуре F обозначается, как обычно, $A \in F$. Точки здесь, конечно, те же, что в принятых выше аксиомах.

Формулируем аксиомы фигуры; их всего три.

1. Основная аксиома фигуры. *Фигура определяется принадлежащими ей точками,* т. е. если фигуры F, F' таковы, что каждая точка, принадлежащая F, принадлежит также F', и обратно: каждая

точка, принадлежащая F' принадлежит также F, то F и F' — это одна и та же фигура.

Эта аксиома не только соответствует представлению о фигурах, но выражает то, как мы судим о совпадении или различии фигур, определенных разными способами. Мы проверяем, является ли любая точка одной фигуры также точкой другой, и обратно. Например, допустим, фигура F задана геометрическим условием, и мы нашли ее уравнение в координатах x, y. Мы говорим, что уравнение задает фигуру F, если координаты каждой ее точки удовлетворяют уравнению и, обратно, каждая пара x, y, удовлетворяющая уравнению, служит координатами некоторой точки фигуры F. Совершенно так же проверяется совпадение фигур, заданных разными геометрическими условиями (точно так же отображение одной фигуры на другую определяется по точкам).

2. Аксиома точки. *Точка есть фигура. Она принадлежит сама себе и никакие другие точки ей не принадлежат.* (Это соответствует определению в «Началах» Евклида: «Точка — то, что не имеет частей».)

3. Аксиома задания фигуры. *Для всякого условия, налагаемого на точки, существует фигура, которая содержит все точки, удовлетворяющие этому условию, и не содержит никаких других.* При этом условие, о котором идет речь, должно быть выражено в понятиях, фигурирующих в принятой аксиоматике (иначе оно не имеет смысла в геометрии, поскольку она полностью строится на принятых аксиомах).

Кроме того, разумеется, условие, налагаемое на точки, должно быть вполне определенным, т. е. для каждой точки должно быть однозначно определено (в принципе проверяемо), выполняется это условие для нее или нет.

При нашей аксиоматике то условие, например, что точка принадлежит данному отрезку AB, является определенным — проверяемым. Также проверяемо, например, и то условие, что точка M находится на данном расстоянии от данной точки A, т. е. отрезок AM имеет данную длину. Условие может, конечно, выражаться не прямо через понятия, входящие в аксиомы, а через такие, которые к ним сводятся, — например, как расстояние от точки до данной прямой и т. п.

Могут быть также условия, которые не выполняются ни для какой точки (например, что при данных точках A, B $|AB| > |AM| + |BM|$). Тогда фигура, определяемая таким условием, не содержит точек: она — пустое множество. Это не исключается, и, соответственно, пустое множество тоже считается фигурой — фигурой без точек.

Прямая (AB) определена выше тем, что ей принадлежат все те и только те точки M, каждая из которых принадлежит вместе с A и B одному отрезку. Такое условие допускается аксиомой 3, и, следовательно, прямая есть фигура. Совершенно так же и луч является фигурой.

О точках, принадлежащих фигуре, имея в виду их все, мы будем говорить, что они *образуют* эту фигуру, и что она является множеством этих точек.

Так как по нашей аксиоматике условие, что точка принадлежит данному отрезку AB, «проверяемо», то согласно аксиоме 3 задания фигуры точки отрезка AB образуют фигуру. Вместе с тем, согласно доказанной выше теореме, отрезок определяется принадлежащими ему точками так же, как фигура. Поэтому мы можем принять, что отрезок AB *является* фигурой, образованной его точками. То есть мы принимаем как определение:

Отрезок является фигурой, образованной его точками. Отношение: точка принадлежит отрезку и точка принадлежит фигуре в этом случае понимается как одно и то же.

Аксиомы фигуры позволяют принять формальное определение:

Фигура — это то, что под этим названием удовлетворяет аксиоме фигуры.

Говоря «множество точек» или «геометрическое место точек», мы имеем в виду то же самое: «множество точек с таким-то условием» — это то же, что «фигура, образованная точками с этим условием». «Образованная» не означает здесь ничего, кроме того, что согласно аксиоме 1 фигура определяется своими точками. Связывать с этим такое представление, что фигура состоит из точек, как вещество из молекул, совершенно необязательно. (В аксиоматике геометрии, данной Гильбертом, прямая мыслится как самостоятельный объект, которому принадлежат точки, но который нив каком смысле из них не состоит.)

§ 6. Плоскостные аксиомы и их первые следствия

Пусть a — какой-нибудь отрезок и $\overline{a}$ — содержащая его прямая.

Мы говорим, что точки C, D лежат *с одной стороны* от прямой $\overline{a}$ и отрезка a, если отрезок CD не имеет с $\overline{a}$ общих точек.

Точки A, B лежат *с разных сторон* от прямой $\overline{a}$ и отрезка a, если отрезок AB пересекает прямую $\overline{a}$ (рис. 5).

III$_1$ (аксиома деления плоскости). *Для каждого отрезка и содержащей его прямой есть не лежащие на ней точки, и все они делятся*

на два класса: точки каждого класса лежат с одной стороны от прямой, точки из разных классов — с разных сторон.

То есть если точки A, B лежат с разных сторон от прямой $\overline{a}$, то каждая не лежащая на ней точка лежит с одной стороны от прямой либо с точкой A, либо с точкой B.

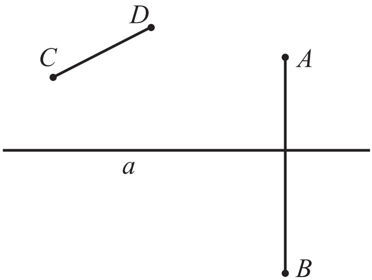


Рис. 5

Полуплоскость. Точки одного класса (из указанных в аксиоме $\mathbf{III}_1$) вместе с точками самой прямой образуют, как принято говорить, *полуплоскость.* Другими словами: фигура, образуемая точками прямой и точками одного класса, называется полуплоскостью, *ограниченной* данной прямой. Приняв это определение, аксиому деления плоскости можно выразить так:

$\mathbf{III}_1$. *Всякая прямая делит плоскость на две полуплоскости. Обе они содержат эту прямую.*

Замечание. Полуплоскость, из которой исключена ограничивающая прямая, называется *открытой*; а когда хотят подчеркнуть, что в полуплоскость включена ограничивающая прямая, то полуплоскость называют *замкнутой.*

Угол. *Углом* мы называем просто пару (несовпадающих) отрезков с общим концом; эти отрезки — *стороны* угла, их общий конец — его *вершина.*

Однако значение имеют только те свойства угла, которые не изменяются, если заменить его стороны на любые отрезки, на них налегающие, с общим концом в той же вершине. О таких парах отрезков мы говорим, что они *представляют* или *образуют один и тот же угол* и также служат его *сторонами.* То есть можно сказать, что угол рассматривается с точностью до «удлинения или укорочения» сторон.

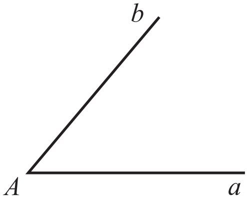


Рис. 6

В этом смысле мы понимаем выражение «данный угол». Он задается двумя сторонами, и если их заменить другими, налегающими на данные, то это будет все тот же «данный угол» и новые стороны будут все так же его сторонами.

Ниже будет доказана

Теорема 1. *Для взаимного расположения двух отрезков с общим концом есть три и*

только три возможности:

(а) *Каждый из отрезков лежит целиком с одной стороны от другого, т. е. все его точки, кроме одного конца, лежат с одной стороны* (рис. 6).

(б) *Отрезки составляют один отрезок* (рис. 7).

(в) *Отрезки налегают друг на друга.*

Рис. 7

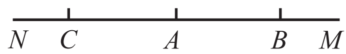


Рис. 8

Соответственно различаются три вида углов: угол с наложенными сторонами — ***нулевой***; в дальнейшем этот случай исключается из рассмотрения. Угол со сторонами, составляющими один отрезок, — ***развернутый.*** Угол, образованный отрезками, лежащими каждый с одной стороны от другого, — ***настоящий***, или, как обычно говорят, ***неразвернутый.***

Доказательство теоремы 1. Пусть даны отрезки AB, AC. Если отрезок AB не имеет (кроме точки A) общих точек ни с каким отрезком, содержащим AC, то он лежит по одну сторону от AC (по самому определению расположения с одной стороны).

Допустим, отрезок AB имеет, помимо A, еще общую точку с отрезком a, содержащим AC. В таком случае по аксиоме соединения отрезков I_4 отрезки AB и a образуют один отрезок MN. Так что AB и AC содержатся в MN. Можно считать, что точка A лежит на MN (если бы A совпадала, например, с M, то можно продолжить отрезок MN за точку M).

Точка A делит отрезок MN на AM и AN, и возможны два случая:

1) AB, AC содержатся в одном из этих отрезков. Тогда один из них (более длинный), как следует из теоремы 4, § 4, содержит другой, т. е. они налегают друг на друга.

2) AB, AC содержатся в разных отрезках AM, AN, например, $AB \subset AM$, $AC \subset AN$ (рис. 8). Тогда точки B, A, C расположены на отрезке MN в порядке возрастания длин отрезков MB, MA, MC (согласно той же теореме 4, § 4). Поэтому точка A лежит на отрезке BC. А это и значит, в силу аксиомы I_3 деления отрезка, что отрезки AB, AC образуют один отрезок BC.

Таким образом, теорема 1 доказана. □

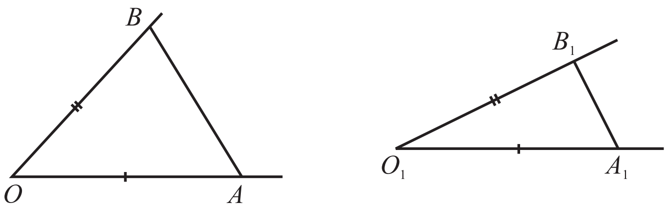
Если отрезок a служит стороной, а его конец A — вершиной неразвернутого угла, то говорим, что угол ***отложен*** от отрезка a, от его

конца A по ту сторону от отрезка a, где лежит другая сторона угла.

Угол обозначается либо вершиной, либо сторонами: $\angle O$ — угол с вершиной O, $\angle ab$ — угол со сторонами a, b; когда же стороны обозначаются концами, то угол со сторонами AB, AC обозначается BAC или CAB (или $\angle BAC$, $\angle CAB$).

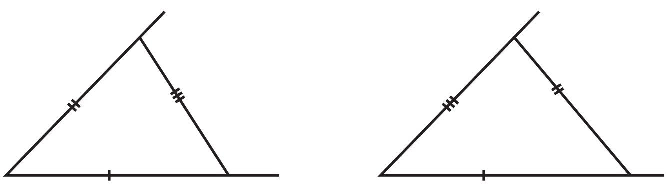
Поперечиной угла назовем отрезок с концами на сторонах угла. Поперечины AB, A_1B_1 углов O, O_1 называем ***соответственными***, если $OA \cong O_1A_1$, $OB \cong O_1B_1$ (рис. 9).

Углы мы называем ***равными***, если у них есть равные соответственные поперечины (рис. 10).



а *б*

Рис. 9



а *б*

Рис. 10

III$_2$ (аксиома равных углов и откладывания угла). *У равных углов все соответственные поперечины равны. От любого данного отрезка от данного его конца и по данную сторону от него можно отложить угол, равный данному неразвернутому углу, и это можно сделать единственным образом.*

Поясним наглядный практический смысл этой аксиомы. Представим себе угол, начерченный на какой-нибудь плоской поверхности; чтобы начертить равный угол в другом месте, можно приложить к

сторонам начерченного угла два стержня, как-то скрепив их в вершине, и перенести их в нужное место. Для того чтобы при перенесении отрезки не «разболтались» и угол не изменился, мы скрепим их какой-нибудь поперечиной (рис. 11, поперечина здесь не отрезок, а реальный стержень). Перенеся угол с поперечиной и начертив по нему угол в нужном месте, мы и проделываем операцию, о которой в отвлеченном виде сказано в аксиоме III_2.

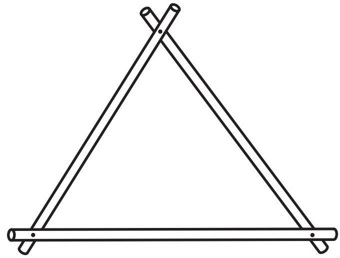


Рис. 11

Введем еще аксиому о мере углов, аналогичную аксиоме II_2. Для этого нам потребуется следующее

Определение. Будем говорить, что отрезок *c проходит внутри угла ab*, если один конец отрезка c находится в вершине угла и отрезок c или его продолжение пересекает какую-нибудь поперечину угла ab. (Тогда луч, идущий из вершины угла вдоль отрезка c, пересекает всякую поперечину угла, как это будет доказано.) Если же угол развернутый, то всякий отрезок, исходящий из его вершины и не налегающий ни на одну из сторон, считается проходящим внутри этого угла.

$\mathbf{III}_3$ (аксиома меры угла). *Если некоторому углу ε отнесено число единица, то каждому углу можно сопоставить положительное число — **численную меру** этого угла **в масштабе** ε, — чтобы выполнялись условия*:

(а) *Равные углы имеют одну и ту же меру.*

(б) *Если отрезок c проходит внутри угла ab, то мера угла ab равна сумме мер углов ac и bc.*

Далее можно было бы принять аксиому существования угла с данной мерой, меньшей меры развернутого угла (аналогично аксиоме $\mathbf{II}_3$ существования отрезка данной длины). Но в этой аксиоме нет надобности, так как во-первых, существование угла с данной мерой может быть доказано, а во-вторых, оно в элементарной геометрии не используется (в ней имеют дело с тригонометрическими функциями, а они определяются через длины отрезков).

Наконец, можно было бы принять для углов аксиому о замене масштаба: при изменении масштаба все численные меры углов умножаются на одно и то же число. Но это можно доказать совершенно так же, как доказывается аналогичное утверждение для отрезков (§ 4, теорема 2). Разница лишь в том, что мера углов ограничена: она не может

превосходить меру развернутого угла. Однако это не изменяет доказательства. Отсюда же следует и единственность численной меры угла.

Из определения равенства углов вытекает важное для их измерения следствие:

Теорема 2. *Все развернутые углы равны.*

Доказательство. Стороны развернутого угла O образуют один отрезок. Поэтому если на них взяты точки A, B, то поперечина AB содержится в этом отрезке и, следовательно, содержит вершину угла O, и, по аксиоме II_2 сложения длин отрезков, $|AB| = |AO| + |BO|$. Поэтому если на сторонах другого развернутого угла O' взяты такие точки A', B', что $O'A' \cong OA$, $O'B' \cong OB$, то $A'B' \cong AB$ (см. теорему 3а, § 4). Это значит, что соответственные поперечины равны, так что углы равны, что и требовалось доказать. □

В геометрии принято измерять углы двумя масштабами: градусами и радианами. По определению, градус 1° — это такой угол, что развернутый угол равен 180°, а радиан — это такой угол, что развернутый угол равен π радианов.

Фундаментальное различие между мерой углов и длиной отрезков состоит в том, что есть геометрически выделенные углы, как развернутый угол или прямой угол, но никаких геометрически выделенных отрезков нет.

Как сторона неразвернутого угла расположена по одну сторону от другой его стороны, так по ту же сторону располагается и всякая поперечина угла. Это непосредственно следует, в силу теоремы 1, из того, что концы поперечины расположены на сторонах угла. На том же основании заключаем, что всякий отрезок с концом в вершине угла, пересекающий поперечину, лежит с той же стороны (рис. 12).

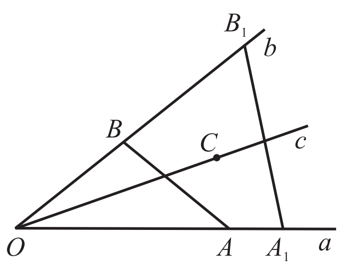


Рис. 12

Теорема 3. *Если отрезок, проведенный из вершины угла, пересекает какую-либо его поперечину, то он или некоторое его продолжение (не за вершину) пересекает любую другую поперечину.*

Доказательство. Пусть имеется угол ab с вершиной O, и пусть отрезок OC пересекает поперечину AB(A на a, B на b; см. рис. 12). Это значит, что точки A, B лежат с разных сторон от отрезка OC, и, стало быть, как следует из теоремы 1, стороны угла ab лежат с разных сторон

от OC. Поэтому если A_1 на a, B_1 на b, то отрезок A_1B_1 пересекает отрезок c, содержащий OC. Вместе с тем отрезки b и OC лежат с одной стороны от a. Поэтому отрезок A_1B_1 пересекает c в точке, лежащей с той же стороны от a, т. е. в точке, принадлежащей отрезку OC или его продолжению за точку C, что и требовалось доказать. □

Треугольником будем называть объединение трех отрезков, соединяющих попарно три точки, не лежащие на одном отрезке. Эти точки — *вершины* треугольника, отрезки — его *стороны*. Треугольник обозначается своими вершинами: ABC (или $\triangle ABC$) и т. п.

Теорема 4. *Если прямая не проходит ни через одну из вершин треугольника и пересекает его сторону, то она пересекает еще одну и только одну его сторону.*

Доказательство. Пусть прямая a пересекает сторону AB треугольника ABC и не проходит через вершину C. Тогда точки A, B лежат по разные стороны от прямой a, и по аксиоме деления плоскости точка C лежит по одну сторону от нее либо с точкой A, либо с B.

В первом случае, когда C и A лежат с одной стороны от прямой a, то C, B — с разных сторон; тем самым прямая a пересекает отрезок BC и не пересекает AC. Во втором случае, напротив, прямая a пересекает AC и не пересекает BC. Теорема доказана. □

Теорема 5. *Если два разных угла ab, ac отложены по одну сторону от отрезка a, то либо отрезок b проходит внутри угла ac, либо c — внутри угла ab.*

Доказательство. Пусть углы ab, ac с вершиной O отложены по одну сторону от отрезка a. Возьмем на сторонах a, b точки A, B и еще точку D на продолжении отрезка a за вершину O. Проведя отрезки AB, BD, AD, получим треугольник ABD (рис. 13).

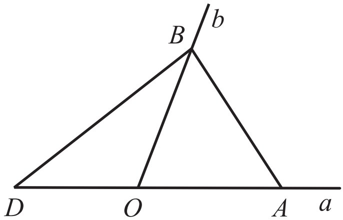


Рис. 13

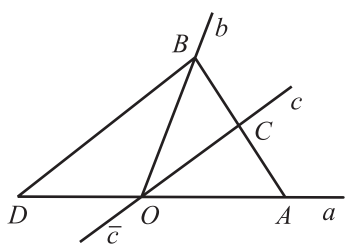


Рис. 14

Прямая $\overline{c}$, содержащая отрезок c, пересекает сторону AD треугольника ABD в точке O и не проходит через вершину B, стало быть, по теореме 4, она пересекает еще одну сторону.

Допустим, она пересекает сторону AB в некоторой точке C (рис. 14). Сторона AB — это поперечина угла ab. Поэтому отрезок OC проходит внутри этого угла. Отрезок c на прямой $\overline{c}$ имеет тот же конец O и расположен с той же стороны от a. Поэтому он налегает на OC и, стало быть, проходит внутри угла ab.

Допустим теперь, что прямая $\overline{c}$ пересекает сторону DB в некоторой точке C' (рис. 15). Проведя отрезок AC', получим треугольник $AC'D$.

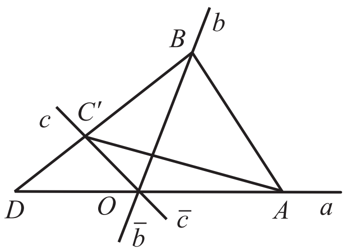


Рис. 15

Прямая $\overline{b}$, содержащая отрезок b, пересекает сторону AD треугольника $AC'D$ и не проходит через вершину. Она поэтому пересекает одну из сторон AC' или DC'. Но сторону DC' она не может пересекать, так как пересекает ее продолжение в точке B. Поэтому $\overline{b}$ пересекает AC', и это значит, что отрезок b проходит внутри угла ac (как это заключаем подобно предыдущему выводу об отрезке c).

Таким образом, выходит, что либо c проходит внутри угла ab, либо b — внутри угла ac, что и требовалось доказать. □

По аксиоме III_3 меры равных углов равны. Вместе с этим выполняется и обратное утверждение.

Теорема 6. *Если меры углов равны, то и сами углы равны.*

Это аналогично тому, что по теореме 3, § 4, отрезки равны, если их длины равны.

Так же, подобно следствию теоремы 3а, § 4, об отрезках выполняется аналогичная теорема об углах.

Теорема 7. *Равенство углов рефлексивно, симметрично и транзитивно.*

Доказательства теорем 6 и 7 совершенно аналогичны доказательствам соответствующих утверждений из § 4 об отрезках. Нужно лишь пересказать те доказательства применительно к углам. Читатель проделает это в качестве упражнения. (В доказательстве теоремы 6 нужно воспользоваться теоремой 5.) □

Смежные углы. Прямые углы. Два угла называются *смежными,* если у них одна сторона общая, а две другие образуют один

отрезок (продолжают друг друга, рис. 16). Эти стороны образуют, стало быть, развернутый угол. Поэтому, в силу аксиомы III_3 меры угла, сумма мер смежных углов равна мере развернутого угла (которая, по теореме 2, одна и та же для всех развернутых углов и равна 180°).

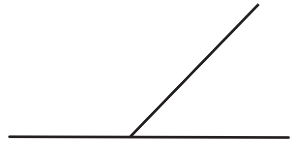


Рис. 16

Отсюда в силу теоремы 6 следует, что углы, смежные равным углам, равны. В частности, *вертикальные* углы, определяемые как смежные с одним и тем же углом, равны.

Угол, равный своему смежному, называется *прямым*. Из предыдущего ясно, что угол, равный прямому, сам прямой.

Теорема 8. *Все прямые углы равны. Мера прямого угла составляет половину меры развернутого угла.*

Докажем это со всеми ссылками. Прямой угол равен своему смежному. Поэтому, в силу аксиомы III_3, их меры равны и, так как сумма этих мер равна мере развернутого угла, то мера прямого угла равна половине меры развернутого угла. Следовательно, меры у всех прямых углов равны. В силу теоремы 6, все прямые углы равны. Что и требовалось доказать. □

Замечание. В «Началах» Евклида равенство прямых углов принято без доказательства, как аксиома. Именно IV постулат Евклида гласит: «Все прямые углы равны».

В § 1 следующей главы мы докажем существование прямых углов.

§ 7. Аксиома параллельных

К плоскостным аксиомам принадлежит в качестве последней аксиомы планиметрии аксиома параллельных. Мы сформулируем ее не для прямых, как это обычно принято, а для отрезков, поскольку отрезки служат у нас основными объектами.

$\mathbf{IV}_1$ (аксиома параллельных отрезков). *Если отрезки AC, BD равны, лежат по одну сторону от отрезка AB и образуют с ним прямые углы, то отрезок CD* равен AB (рис. 17).

Замечание к аксиоме $\mathbf{IV}_1$. Отрезки, обладающие свойствами отрезков AC, BD, фигурирующими в аксиоме, мы называем *параллельными.* Любые равные друг другу отрезки AM, BN, налегающие на отрезки AC, BD, обладают теми же свойствами, и, по аксиоме, $NB \cong AB$. Таким образом, параллельные отрезки можно продолжать

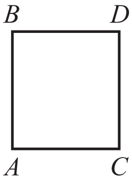


Рис. 17

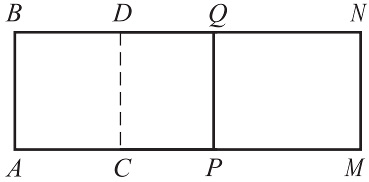


Рис. 18

сколь угодно далеко, и отрезки PQ между их точками P, Q, для которых $AP = BQ$, все равны друг другу (рис. 18). Как говорят, отрезки AM, BN проходят — расположены — *на постоянном расстоянии* друг от друга, как рельсы прямой линии железной дороги (рис. 19). «Параллельный» по-гречески и означает «идущий рядом».

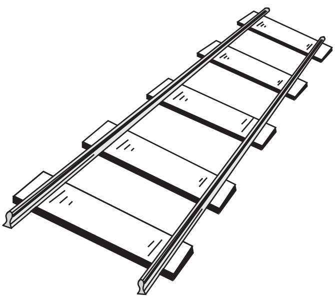


Рис. 19

Легко доказать — и это будет сделано в § 2 гл. II, — что отрезок CD в аксиоме образует с AC и BD прямые углы (и так же любой отрезок PQ образует прямые углы с AP, BQ). Фигура $ABCD$ представляет собой прямоугольник. Другими словами, аксиома IV_1 выражает тот хорошо известный из практики факт, что можно строить прямоугольники с любыми сторонами. Прямоугольниками являются наиболее часто встречающиеся в быту фигуры: полы, стены, окна, стекла, листы бумаги и т. д., и т. п.

Обычно принимают аксиому, говорящую о параллельных прямых. Напомним, что в планиметрии две прямые называются *параллельными,* если они не имеют общих точек; в этом случае также говорят, что одна прямая *параллельна* другой. Для параллельных прямых a, b применяется обозначение: $a||b$ или $b||a$.

Аксиома параллельных прямых. *Для каждой прямой a и не лежащей на ней точки A существует не более одной прямой, параллельной данной прямой a и проходящей через точку A.*

То, что через данную точку A, не лежащую на данной прямой a, проходит по крайней мере одна прямая, параллельная a, легко доказывается (это прямая, перпендикулярная перпендикуляру AB к прямой a, как доказано дальше в § 2 гл. II). Аксиома утверждает, что такая прямая только одна.

Здесь говорится о прямых во всем их бесконечном протяжении, поэтому аксиома не имеет прямого практического содержания, и для наглядного ее понимания требуется представить себе бесконечные прямые.

Обходясь без представления о прямых, взятых во всем их бесконечном протяжении, имеют в виду конечные, но неограниченно продолжаемые прямые (отрезки) и говорят: прямые параллельны, если они не пересекаются, как их ни продолжать. (Когда говорят о параллельных линиях в практике, то имеют в виду, что они проходят на постоянном расстоянии (о чем сказано выше), но едва ли, что они при бесконечном продолжении не пересекутся. В геометрии же пользуются признаками параллельности, в которых речь идет, на самом деле, об отрезках и углах.) Аксиома параллельных прямых проще по формулировке, но дальше от практики, чем аксиома параллельных отрезков.

Представление о бесконечной прямой было чуждо грекам; у Евклида роль аксиомы параллельных играл пятый постулат, утверждающий: «*Если из концов данного отрезка проведены в одну сторону от него два отрезка, образующие с ним углы, в сумме меньшие развернутого угла, то сами эти отрезки или их продолжения пересекаются* (имеется в виду продолжение в ту же сторону от данного отрезка, см. рис. 20)[6].

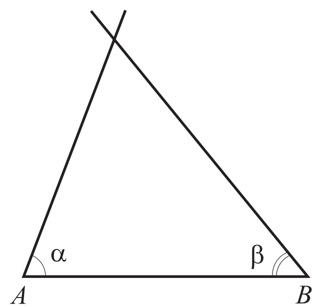


Рис. 20

Это утверждение равносильно следующему:

При всяком данном отрезке AB и данных двух углах α, β, в сумме меньших развернутого, существует (можно построить) треугольник со стороной AB и прилежащими углами $A = \alpha$ и $B = \beta$.

Отрезки, проведенные в одну сторону от отрезка AB под данными углами, согласно V постулату пересекутся, так что и получится нужный треугольник. Поэтому V постулат имеет практический смысл условия, при котором можно построить треугольник с двумя данными углами.

[6]Буквально формулировка в «Началах» Евклида несколько иная, но в точности равносильная. В других списках «Начал» то же утверждение фигурирует как XI аксиома.

Все три аксиомы: параллельных отрезков, параллельных прямых, пятый постулат — равносильны. Это мы докажем позже, в § 2, гл. II.

Замечание. Сложность V постулата и то, что он вовсе не так очевиден, как остальные, вызвала попытки вывести его из других предпосылок геометрии — попытки, длившиеся со времен Евклида свыше 2000 лет, пока не было осознано, что такой вывод невозможен. Геометрия, в которой этот постулат отрицается, была построена Н. И. Лобачевским (1792–1856) и Я. Бояи (1802–1860) в 20-х годах XIX в.

Учителю полезно иметь в виду разные формы аксиомы параллельных — хотя бы для возможности выбора.

§ 8. Аксиомы стереометрии и их первые следствия

В аксиоматике стереометрии наряду с основными понятиями планиметрии вводятся новые о б ъ е к т ы — *плоскости* и о т н о ш е н и е — точка *принадлежит* плоскости или, что то же, плоскость *проходит* через точку. Вводится определение: отрезок *содержится* (*лежит*) *в* плоскости, если все его точки принадлежат этой плоскости; так же определяется, что прямая содержится в плоскости.

Аксиомы образуют три группы: 1) линейные, 2) плоскостные, 3) пространственные.

Линейные аксиомы дословно повторяют линейные аксиомы § 3, 4.

Плоскостные аксиомы повторяют плоскостные аксиомы § 6, 7 с той разницей, что теперь к трем из них добавляется указание, что они относятся к каждой плоскости. Именно:

Аксиома деления плоскости. *На каждой плоскости по отношению ко всякому содержащемуся в ней отрезку различаются две стороны.*

Аксиома откладывания угла. *На каждой плоскости по данную сторону от лежащего в ней отрезка от данного его конца можно отложить угол, равный данному.*

Аксиома параллельных. *На каждой плоскости...* и далее формулируется аксиома параллельных отрезков или прямых.

Формулировки остальных аксиом можно не менять.

Пространственные аксиомы; их всего три.

Пр. 1. Через каждые три точки проходит плоскость.

Пр. 2. Существуют четыре точки, не лежащие в одной плоскости.

Пр. 3. Если две плоскости имеют общую точку, то их пересечение представляет собой прямую. (Прямая определяется через отрезки, как это сделано в § 5.)

Так как линейные аксиомы тут те же, что в планиметрии, то можно воспользоваться всеми выводами из них, полученными раньше. В частности, имеет место

Теорема 1. *Через каждые две точки проходит прямая, и притом только одна.* □

Теперь докажем следующую теорему.

Теорема 2. *Если прямая имеет с плоскостью две общие точки, то она содержится в этой плоскости.*

Доказательство. Пусть две точки A, B прямой a лежат в плоскости α. К точкам A, B можно присоединить какую-нибудь точку C, не лежащую в α (как это следует из аксиомы Пр. 2). Через точки A, B, C проходит плоскость β (аксиома Пр. 1). По аксиоме Пр. 3 плоскости α и β пересекаются по прямой. Эта прямая содержит точки A, B. Но по теореме 1 через две точки проходит только одна прямая. Это, стало быть, и есть данная прямая a, и она лежит в плоскости α. Что и требовалось доказать. □

Из теорем 1, 2 очевидно вытекает

Следствие. *Во всякой плоскости каждые две точки можно соединить отрезком, и притом только одним.* □

Вместе с линейными и плоскостными аксиомами это приводит к выводу:

Теорема 3. *На каждой плоскости выполняется планиметрия.* □

Установим еще два свойства плоскости как фигуры в пространстве.

Теорема 4. *Через каждые три точки, не лежащие на одной прямой, проходит плоскость, и притом только одна.*

Доказательство. Пусть A, B, C — точки, не лежащие на одной прямой. Через них, по аксиоме Пр. 1, проходит плоскость. Эта плоскость α только одна, потому что всякая другая плоскость β, имеющая с α общие точки, пересекает α по прямой (аксиома Пр. 3) и, стало быть, не может иметь с α общих точек, не лежащих на одной прямой. Теорема 4 доказана. □

Следствие 1. *Через прямую и не лежащую на ней точку проходит плоскость, и притом только одна.*

Следствие 2. *Через две пересекающиеся прямые проходит плоскость, и притом только одна.*

Оба следствия выводятся из теоремы 4 с использованием теоремы 2. Читатель докажет их самостоятельно. □

Теорема 5. *Всякая плоскость α делит все не принадлежащие ей точки на два таких класса, что отрезок, соединяющий точки одного класса, не пересекает α, а отрезок, соединяющий точки разных классов, пересекает α.*

Доказательство. Берем какую-нибудь точку A, не принадлежащую α, и относим в класс F_1, вместе с точкой A, все такие точки M, что отрезок AM не имеет с α общих точек; все остальные точки, не принадлежащие α, относим в класс F_2.

Пусть M, N — две точки, отличные от A и не лежащие в плоскости α. Через точки A, M, N проходит плоскость β (аксиома Пр. 1). Отрезки AM, AN, MN лежат в β (следствие теорем 1, 2). Если плоскость β не имеет с α общих точек, то, очевидно, $M,\ N \in F_1$, и отрезок MN не имеет с α общих точек.

Допустим, плоскости α и β имеют общую точку. Тогда, в силу аксиомы Пр. 3, они пересекаются по прямой. Эта прямая a разбивает плоскость β за вычетом самой прямой a на две открытые полуплоскости (в силу аксиомы деления плоскости). Из определения классов F_1, F_2 ясно, что точки одной из этих полуплоскостей входят в F_1, точки другой — в F_2. По свойству полуплоскостей, если точки M, N принадлежат одной из них, то отрезок MN не пересекает прямую a. Если же точки M, N принадлежат разным полуплоскостям, то отрезок MN пересекает прямую a. А так как точки полуплоскостей содержатся в классах F_1 и F_2, то получаем сказанное в теореме: MN не пересекает α, если точки M, N из одного класса, и пересекает, если точки из разных классов; что и требовалось доказать. $\square$

Дополнение. Другой вариант аксиом стереометрии. Считая планиметрию известной, можно, переходя к стереометрии, принять, что на каждой плоскости выполняется планиметрия. Подобная аксиоматика удобна при таком изложении геометрии, когда сначала излагается планиметрия, а потом — стереометрия (при этом безразлично, на каких аксиомах строится сама планиметрия).

Такое изложение дано в книге: *Александров А. Д., Вернер А. Л., Рыжик В. И.* Геометрия. 9–10: Учебное пособие для школ и классов с углубленным изучением математики. М.: Просвещение, 1984. Там используется понятие величины — расстояния между точками, но его можно заменить численным расстоянием. Тогда аксиомы стереометрии представляются следующим образом (с теми же основными объектами и отношениями):

1. (Аксиома расстояния.) Каждым двум точкам сопоставлено расстояние — положительное число (или величина).

2. (Аксиома плоскости.) Существует хотя бы одна плоскость, и на каждой плоскости выполняется планиметрия. При этом численная длина каждого отрезка равна данному по аксиоме 1 расстоянию между его концами.

Дальше следуют три пространственные аксиомы — те же, что и выше: Пр. 1, Пр. 2, Пр. 3.

Глава II

НАЧАЛА ЭЛЕМЕНТАРНОЙ ГЕОМЕТРИИ

§ 1. Треугольники, перпендикуляры

Напомним, что треугольником мы называем фигуру, образованную тремя отрезками с попарно общими концами, не содержащимися в одном отрезке[7]. Стороны, вершины, углы треугольника определяются как обычно, так же как равенство треугольников — по равенству сторон и соответственных углов. Из того, что вершины не содержатся в одном отрезке, следует, что углы треугольника настоящие, и каждая сторона служит поперечиной противоположного угла (см. теорему 1, § 6, гл. 1).

Выполняются три классические теоремы о равенстве треугольников (только мы устанавливаем их в ином порядке):

(I) *по трем сторонам,*

(II) *по двум сторонам и углу между ними,*

(III) *по стороне и двум прилежащим углам.*

Доказательство (I). Пусть у треугольников ABC, $A_1B_1C_1$ равны стороны: $AB = A_1B_1$ и т. д. Ввиду этих равенств отрезки BC, B_1C_1 являются соответственными поперечинами углов A, A_1. По условию $BC = B_1C_1$. Значит, $\angle A = \angle A_1$. Аналогично доказывается, что $\angle B = \angle B_1$ и $\angle C = \angle C_1$ (рис. 21). $\square$

Доказательство (II). Пусть у треугольников ABC, $A_1B_1C_1$

$$AB = A_1B_1, \quad AC = A_1C_1, \quad \angle A = \angle A_1.$$

Ввиду первых двух равенств у углов A, A_1 поперечины BC, B_1C_1 соответственные, и по равенству углов они равны. Поэтому из (I) треугольники равны (рис. 22). $\square$

---

[7] Треугольником называют и каркас, и часть плоскости, как в словах: «треугольник из проволоки» или «вырежьте треугольник из бумаги», и три точки, не лежащие на одной прямой, и три соединяющие их отрезка.

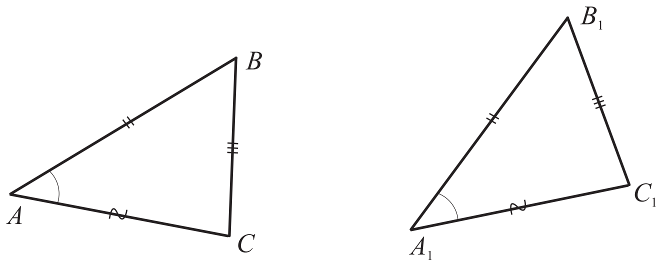


Рис. 21

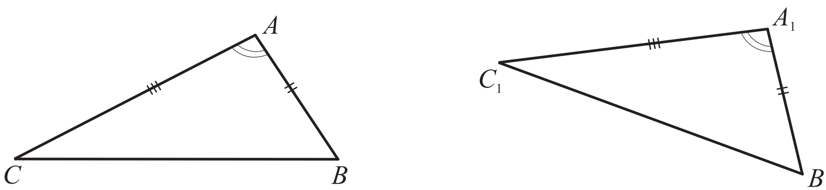


Рис. 22

Доказательство (III). Пусть у треугольников ABC, $A_1B_1C_1$

$$AB = A_1B_1, \quad \angle A = \angle A_1, \quad \angle B = \angle B_1.$$

Отложим вдоль A_1C_1 отрезок A_1C_2, равный AC. Тогда ввиду (II) треугольник $A_1B_1C_2$ равен треугольнику $\triangle ABC$. Стало быть, его угол B_1 равен углу B в треугольнике ABC. По аксиоме откладывания угла, угол, равный B, может быть здесь только один, так что угол B_1 тут один и тот же (в $\triangle A_1B_1C_1$ и $\triangle A_1B_1C_2$). Тем самым $\triangle A_1B_1C_2$ совпадает с $\triangle A_1B_1C_1$, и, значит, $\triangle A_1B_1C_1 = \triangle ABC$ (рис. 23). □

Перпендикуляры. Мы говорим, что отрезок a *перпендикулярен* отрезку b, если сами эти отрезки или их продолжения пересекаются под прямым углом. В записи: $a \perp b$.

Теорема 1. *Для всякого отрезка a и любой точки A через нее проходит отрезок, перпендикулярный a.*

Доказательство. Принципиально различаются два случая: а) точка A лежит на каком-нибудь отрезке, содержащем a; б) это не так. Нам будет удобнее начать со случая б).

Случай б). Возьмем на a какую-либо точку B, и пусть BC — один из отрезков, на которые B делит a. Проведем отрезок AB; возьмем

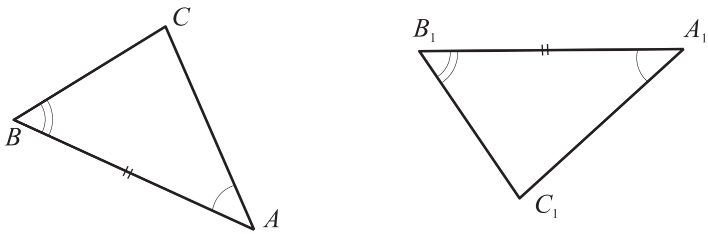


Рис. 23

угол ABC и отложим от BC равный ему угол A_1BC с другой стороны от отрезка BC. На его стороне мы можем взять точку A_1 с $A_1B = AB$ (рис. 24).

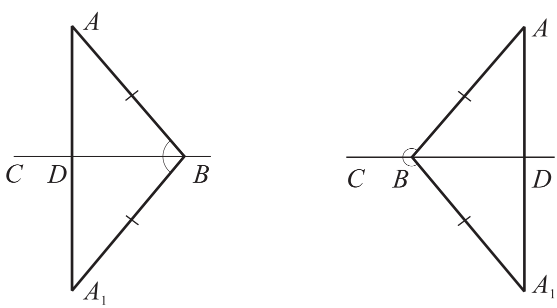


Рис. 24

Точки A и A_1 лежат с разных сторон от a, поэтому отрезок AA_1 пересекает некоторый отрезок a', содержащий a, в какой-то точке D. Может случиться, что D совпадает с точкой B, так что отрезки AB и A_1B образуют один отрезок AA_1. Тогда углы, образуемые ими с отрезком BC, смежные, а так как по построению они равны, то они прямые, так что $AB \perp a$.

Пусть точка D отлична от B, так что имеются два треугольника ABD и A_1BD. У них $BA = BA_1$, сторона BD общая (и так как отрезок равен самому себе, то стороны BD равны). Кроме того, углы при вершине B равны либо по построению, либо как смежные построенным равным углам ABC и A_1BC. Поэтому треугольники ABD и A_1BD равны, и, стало быть, равны их углы D. А так как они смежные, то значит прямые. Таким образом, $AA_1 \perp a$.

С л у ч а й а). Пусть для простоты, точка A лежит на отрезке a, так что имеем на a отрезки AB, AB_1. Пользуясь только что изложенным построением, построим где-нибудь прямой угол d. Отложим равный ему угол BAC от отрезка AB, от его конца A. Он прямой, поскольку равен прямому углу. Следовательно, $AC \perp a$. □

Это построение как бы воспроизводит откладывание прямого угла с помощью угольника.

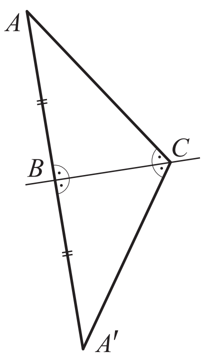


Рис. 25

Для доказательства единственности перпендикуляра нам нужна

Лемма (обратная теореме о сумме смежных углов). *Если у двух углов одна сторона общая, а две другие расположены по разные стороны от нее, и в сумме углы составляют* 180°, *то они смежные.* Докажите это самостоятельно. □

Теорема 1. *Перпендикуляр, проведенный из данной точки к данному отрезку, единственный.*

Д о к а з а т е л ь с т в о. Допустим противное, и пусть отрезки AB, AC перпендикулярны BC. Продолжим отрезок AB по другую сторону от BC на отрезок $BA' = AB$. Треугольник $A'BC$ равен $\triangle ABC$. Поэтому их углы C равны. Применяя предыдущую лемму, приходим к противоречию (рис. 25). □

§ 2. Параллельность. Метрические соотношения в треугольнике

В связи с аксиомой IV_1 «О параллельных отрезках» мы назвали параллельными равные отрезки AC, BD, проведенные в одну сторону из концов какого-либо данного отрезка AB перпендикулярно к нему. Аксиома утверждает, что в таком случае $CD = AB$ (рис. 17).

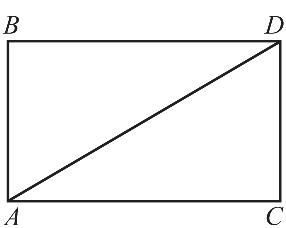


Рис. 26

Лемма 1. *Параллельные отрезки перпендикулярны также отрезку CD.*

Д о к а з а т е л ь с т в о. Проведя отрезок AD, получим два треугольника ABD и ACD (рис. 26). У них $AC = BD$ по условию, $AB = CD$ по аксиоме, сторона AD — общая. Поэтому углы этих

треугольников равны, в частности $\angle B = \angle C$. Но угол B по условию прямой. Следовательно, и угол C прямой, так что $AC \perp CD$. Для BD вывод тот же. □

Таким образом, здесь все отрезки играют попарно одинаковую роль: противоположные равны и в общих концах взаимно перпендикулярны; короче, они образуют прямоугольник.

Сумма углов треугольника. Вот еще одна переформулировка V постулата.

Лемма 2. *Сумма углов прямоугольного треугольника равна развернутому углу.*

Доказательство. Пусть ABC — треугольник с прямым углом A. Проведем отрезок BD, равный AC, в ту же сторону от AB под прямым углом к AB (рис. 27). Получаем прямоугольник. Треугольники ABC и DBC равны по трем сторонам. Так как они составляют прямоугольник, то сумма всех их углов составляет четыре прямых угла. А значит, у одного треугольника сумма углов равна двум прямым. □

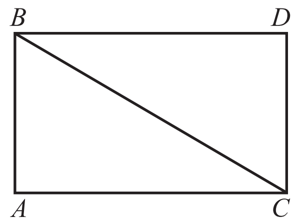


Рис. 27

Из доказанного очевидно вытекает

Следствие. *У прямоугольного треугольника два угла острые, и их сумма равна прямому углу.* □

Лемма 3. *Во всяком треугольнике два угла острые. Если в треугольнике ABC углы A, B острые, то основание высоты CH лежит на AB.*

Доказательство. Если в треугольнике один угол прямой, то два другие острые (предыдущее следствие). Допустим, в треугольнике ABC угол B тупой. Проведем высоту CH. Если бы точка H лежала на луче BA, то у прямоугольного треугольника CBH угол B был бы тупой (рис. 28), что невозможно. Значит, H лежит с другой стороны от точки B, и мы имеем прямоугольный треугольник ACH, причем B на AH. Его углы A, C острые. Значит, и углы A, C треугольника ABC острые (рис. 29).

Пусть у треугольника ABC углы A, B острые. Если бы основание H высоты CH не лежало на AB, то мы имели бы прямоугольный треугольник с тупым углов — или $\triangle ACH$ с углом A (рис. 30), или $\triangle HCB$ с углом B. Но ни то, ни другое невозможно. Лемма доказана. □

Будем обозначать прямой угол (точнее, его величину) буквой d.

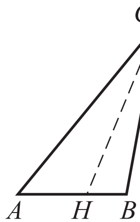


Рис. 28

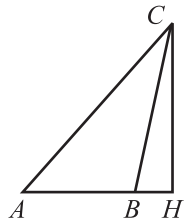


Рис. 29

Теорема 1. *Сумма углов любого треугольника равна* $2d$, *т. е. развернутому углу.*

Доказательство. Пусть дан треугольник ABC. По лемме 3 у него два угла острые, пусть это будут углы A, B. Тогда по той же лемме 3 основание H высоты CH лежит на AB. Поэтому имеем два прямоугольных треугольника $\triangle ACH$ и $\triangle BCH$ (рис. 30). Суммы их углов равны $2d$, так что у обоих вместе — $4d$. Но углы при H дают $2d$. Поэтому сумма остальных углов равна $2d$:

$$\angle A + \angle B + \angle C_1 + \angle C_2 = \\ = \angle A + \angle B + \angle C = 2d,$$

что и требовалось доказать. □

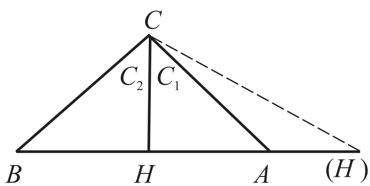


Рис. 30

Углы при параллелях и секущей. Определим *параллельные отрезки* a, b (короче — *параллели*) как такие, у продолжений которых $\overline{a}$, $\overline{b}$ есть общий перпендикуляр (т. е. такой отрезок AB, что $A \in \overline{a}$, $B \in \overline{b}$, $AB \perp a$ и $AB \perp b$). Доказанная в начале параграфа лемма 1 равносильна следующему.

Лемма 1а. *Если вдоль параллельных отрезков* a, b *от концов их общего перпендикуляра* AB *отложить в одну сторону равные отрезки* AC, BD, *то* CD — *также их общий перпендикуляр. При этом он равен* AB (см. рис. 26). □

Сказанное можно выразить еще так.

Из каждой точки одного из параллельных отрезков идет ко второму или его продолжению их общий перпендикуляр. (Здесь второй отрезок представляется, если нужно, продолженным так, чтобы на нем самом был конец общего перпендикуляра.)

Отрезок, пересекающий две параллели — параллельные отрезки, — называется *секущей.* Углы с вершинами в точках пересечения A, B

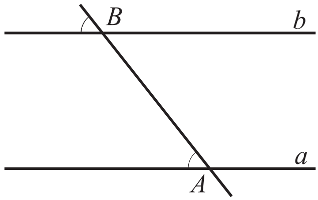


Рис. 31

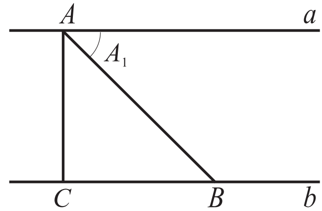


Рис. 32

объединяются, как обычно, в пары, такие как внутренние накрест лежащие и т. д. Им можно дать формальные определения. *Внутренние* — это углы, образуемые отрезком AB с отрезками, на которые параллели делятся точками A и B. *Накрест лежащие* — это те, стороны которых лежат по разные стороны от AB, а вершины — у одного A, у другого B. *Соответственные* — это такие два угла, один из которых внутренний, а другой вертикальный накрест лежащему (рис. 31).

Теорема 2. *Внутренние накрест лежащие углы при параллелях равны.*

Доказательство. Пусть точки A, B лежат на параллельных отрезках a, b: A на a, B на b. Проведем отрезок AB и из точки A — перпендикуляр AC к отрезку b. В силу леммы 2 он также перпендикулярен a (рис. 32). Треугольник ABC прямоугольный, поэтому в нем $\angle A + \angle B = d$. А так как $AB \perp a$, то углы при A с общей стороной AB составляют прямой угол, т. е. $\angle A + \angle A_1 = d$. Вместе с предыдущим равенством это дает $\angle B = \angle A_1$, что и требовалось доказать. $\square$

Отсюда ввиду определения соответственных углов непосредственно вытекает

Следствие. *Соответственные углы при параллелях равны.* $\square$

Начала тригонометрии. Тригонометрию мы начнем с отношения проекции наклонной к самой наклонной, т. е. с косинуса острого угла.

Наклонная и проекция определяются как обычно. *Наклонная* к отрезку a — это отрезок, не перпендикулярный a, с концом, принадлежащим a. *Проекция отрезка* AB на прямую $\overline{a}$ — это такой отрезок $A'B'$, содержащийся в $\overline{a}$, что $AA' \perp a$ и $BB' \perp a$ (рис. 33) (либо A' совпадает с A, либо B' — с B).

Отношение отрезков мы определим как отношение их длин, измеренных в произвольном масштабе. Из теоремы 5, § 4 о замене масштаба следует, что указанное отношение не зависит от того, в каком

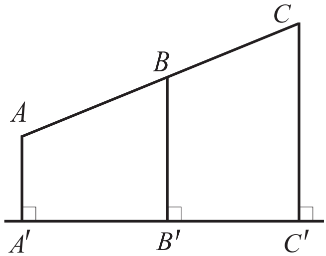


Рис. 33

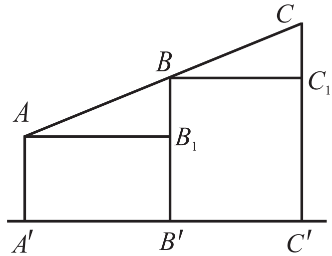


Рис. 34

масштабе выражены длины отрезков. Поэтому его можно записывать как отношение самих отрезков: a/b.

Лемма 4. *Пусть AB, BC — отрезки, составляющие один отрезок AC, и $A'B'$, $B'C'$ — их проекции на один и тот же отрезок. Тогда если $AB = BC$, то $A'B' = B'C'$* (рис. 33).

Доказательство. Пусть выполнены условия леммы и $AB = BC$. Отрезки AA', BB' перпендикулярны одному отрезку, а значит, имеют общий перпендикуляр $A'B'$. Поэтому согласно лемме 1а у них (или у содержащих их отрезков) есть общий перпендикуляр AB_1, причем $AB_1 = A'B'$. Если $B_1 = B$, то все просто. Если $B_1 \neq B$, мы получаем треугольник AB_1B с прямым углом B_1 и с $AB_1 = A'B'$ (рис. 34).

Совершенно так же можно построить треугольник BC_1C с прямым углом C и с $BC_1 = B'C'$.

У этих треугольников, по условию, равны гипотенузы: $AB = BC$. Кроме того, их углы B, C равны как соответственные (поскольку отрезки BB' и CC' параллельны; рис. 34). А так как углы B_1, C_1 прямые, то и другие их углы — A, B — равны (как следует из того, что сумма всех углов одна и та же). Следовательно, треугольники равны, так что $AB_1 = BC_1$. А так как $AB_1 = A'B'$ и $BC_1 = B'C'$, то $A'B' = B'C'$, что и требовалось доказать. □

Замечание. Утверждение, что углы B и C в рассматриваемых треугольниках ABB_1, BCC_1 соответственные, основано на том, что эти треугольники лежат с одной стороны от AC (не считая, конечно, самого отрезка AC). Это так, потому что угол B_1BC_1 прямой и, стало быть, меньше угла B_1BC, так что отрезок BC_1 проходит между BB_1 и BC.

Теорема 3. *Отношение проекции к наклонной одно и то же для наклонных и проекций, образующих равные углы.*

Докажем сначала частный случай этой теоремы.

Теорема 3а. *Отношение проекции к наклонной одно и то же для наклонных и проекций, образующих один и тот же угол, когда наклонные, как и проекции, налегают друг на друга.*

Доказательство. Пусть отрезок a — наклонная к некоторому отрезку a', точка A — ее конец на a'. Возьмем на a какой-нибудь отрезок e, и пусть e' — его проекция на a'. Будем измерять отрезки AB, налегающие на a, масштабом e, а отрезки AM', налегающие на a', — масштабом e' (рис. 35).

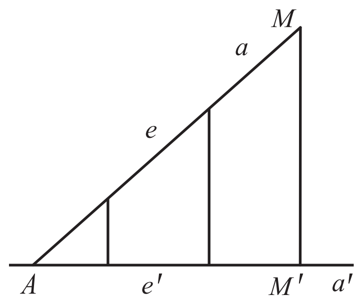


Рис. 35

Пусть отрезок AM составлен из некоторого числа n-х долей отрезка e. Тогда, как следует из леммы 4, его проекция AM' составлена из такого же числа n-х долей отрезка e'. Поэтому длины отрезков AM, AM' в масштабах e, e' численно равны. Так как длина любого отрезка является пределом длин отрезков, составленных из целых долей масштаба, то тот же вывод верен для любого отрезка c наклонной a и его проекции c'. Обозначая длины в масштабах e и e' через l и l', это можно записать так:

$$l'(c') = l(c).$$

По правилу замены масштаба $l'(c') = l(c')/l(e')$. Поэтому

$$\frac{c'}{c} = \frac{l(c')}{l(c)} = l(e').$$

Это и значит, что отношение проекции c' к наклонной c одно и то же для любой наклонной c. Теорема 3а доказана. □

Докажем теперь теорему 3 в общем случае. Пусть имеются две наклонные, образующие со своими проекциями равные углы с вершинами A, B. Отложим вдоль них от точек A, B равные отрезки a, b, и пусть a', b' — их проекции (рис. 36). Отрезки a, b служат гипотенузами, а a', b' — катетами двух прямоугольных треугольников с углами A, B. Так как $a = b$ и $\angle A = \angle B$, то треугольники равны и, стало быть, $a' = b'$. Поэтому $a'/a = b'/b$. По доказанному выше, для любой наклонной c и ее проекции c', налегающих на a и a' имеет место равенство

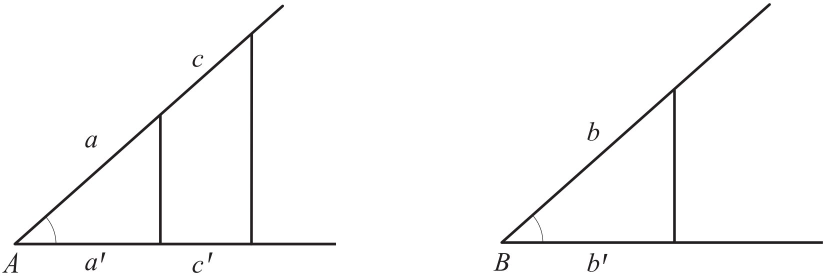


Рис. 36

$c'/c = a'/a$. Следовательно,

$$\frac{c'}{c} = \frac{b'}{b}.$$

Этим равенством и выражается сказанное в теореме. □

Теорема Пифагора. Доказанная теорема 3 представляет основание для вывода всех соотношений между сторонами и углами треугольников, прежде всего — теоремы Пифагора (рис. 37): *квадрат гипотенузы равен сумме квадратов катетов.*

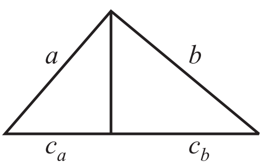


Рис. 37

Пусть a, b — катеты, c — гипотенуза, c_a, c_b — проекции катетов на гипотенузу. По лемме 4 основание высоты лежит на гипотенузе, и поэтому $c_a + c_b = c$. Из теоремы 3 заключаем, что $c_a/a = a/c$, так что $a^2 = c \cdot c_a$. Аналогично, $b^2 = cc_b$. И, складывая, получаем $a^2 + b^2 = c^2$. □

Из теоремы Пифагора очевидно следует

Теорема 4. *Проекция меньше наклонной.* □

Отсюда выводится:

Теорема 5. *Сумма двух сторон треугольника больше третьей.* (Так как эта третья сторона есть либо сумма, либо разность проекций двух других сторон; рис. 38 *а, б.*) □

Тригонометрические функции. Теоремы о треугольниках. Теорема 3 равносильна тому, что отношение проекции и наклонной зависит только от угла между ними, т. е. является его функцией. Это — *косинус* острого угла A (угол между наклонной и проекцией острый), обозначается $\cos A$. Косинус смежного с ним тупого угла определяет-

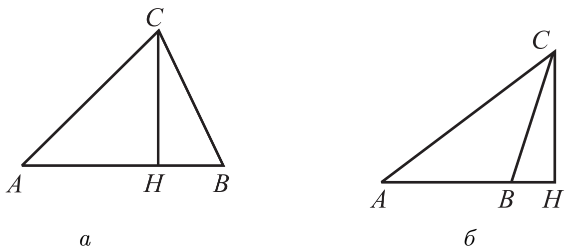


Рис. 38

ся как то же отношение с минусом. Косинус прямого угла полагают равным нулю.

Отношение перпендикуляра, опущенного из конца наклонной, к наклонной — то же, что отношение противолежащего катета к гипотенузе, — зависит только от угла, так как из теоремы Пифагора следует, что

$$\left(\frac{a}{c}\right)^2 = 1 - \left(\frac{b}{c}\right)^2 = 1 - \cos^2 A.$$

Отношение a/c есть *синус* угла A, обозначается $\sin A$. Синус смежного с ним угла определяется тем же отношением.

Далее устанавливаются две основные теоремы о треугольниках: теорема синусов и обобщенная теорема Пифагора, которую обычно называют теоремой косинусов.

Теорема синусов. *Во всяком треугольнике стороны относятся как синусы противолежащих углов:*

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}.$$

Доказательство. Для высоты, опущенной из вершины C, по самому определению синуса, имеем: $h = a \sin B = b \sin A$. Отсюда

$$\frac{\sin A}{a} = \frac{\sin B}{b}.$$

Аналогично получаем равенство

$$\frac{\sin A}{a} = \frac{\sin C}{c}. \quad \square$$

Обобщенная теорема Пифагора (теорема косинусов). *Во всяком треугольнике*

$$c^2 = a^2 + b^2 - 2ab \cos C.$$

Доказательство. Пользуясь теоремой Пифагора и определением косинуса, получаем (рис. 39):

$$c^2 = h^2 + a_c^2, \quad h^2 = b^2 - a_b^2,$$
$$a_b = b\cos C, \quad a_b + a_c = a.$$

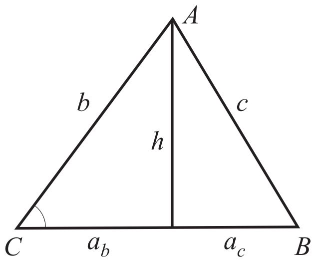


Рис. 39

Отсюда выводится утверждение теоремы. (В первых двух равенствах знаки чисел a_c, a_b не имеют значения!) □

Подведем итог.

Основные соотношения в треугольнике.

1. Сумма углов равна 180°.

2. Теорема синусов (тройное равенство).

3. ОТП — обобщенная теорема Пифагора (три равенства).

4. Выражение площади (три равенства).

С помощью этих теорем можно получить, в принципе, все теоремы о треугольниках. Например, теоремы о подобии.

1. *Если у двух треугольников углы равны, то стороны пропорциональны.* Непосредственно выводится из теоремы синусов.

2. *Если стороны пропорциональны, то углы равны.*

Из ОТП следует: если стороны пропорциональны, то косинусы углов равны, а значит, и углы равны.

3. *Если две пары сторон пропорциональны и углы между ними равны, то все стороны пропорциональны и углы равны.*

Из ОТП следует, что при условиях этой теоремы третьи стороны тоже пропорциональны, и тогда равенство углов следует из 2.

Читатель детально проведет все эти выводы и также проведет выводы других теорем о треугольниках, как указано далее в задачах.

Задачи.

1. Обратная теорема Пифагора. Если $c^2 = a^2 + b^2$, то $\angle C = 90°$.

2. Если две стороны одного треугольника равны двум сторонам другого треугольника, а углы между ними не равны, то третья сторона там больше, где угол больше, и обратно: угол больше, где сторона больше.

3. Во всяком треугольнике сторона меньше суммы и больше разности двух других, например: $a + b > c > |a - b|$.

4. В треугольнике против большей стороны лежит больший угол, и обратно. (С помощью теоремы синусов.)

5. Теорема синусов следует из ОТП. Именно,

$$\sin C = \frac{c\sqrt{2(a^2b^2 + b^2c^2 + c^2a^2) - (a^4 + b^4 + c^4)}}{2abc},$$

так что отношения $\sin C/c$, $\sin A/a$, и $\sin B/b$ равны друг другу.

6. Во всяком треугольнике

$$\cos B - \cos C = \frac{a+b+c}{2abc}(c+b-a)(c-b).$$

Отсюда следует, что против большей стороны лежит больший угол, и обратно.

7. Площадь треугольника выражается формулой Герона (в ней $p = (a+b+c)/2$):

$$S = \sqrt{p(p-a)(p-b)(p-c)}.$$

Выводится из того, что $S = \frac{1}{2}ab\,\sin C$ и результата задачи 5.

8. Пусть в $\triangle ABC$ точка D на AB. Тогда если $AD = c_1$, $BD = c_2$ и $CD = d$, то

$$d^2 = \frac{a^2c_1^2 + b^2c_2^2}{c^2} - c_1c_2.$$

Найдите отсюда медиану и биссектрису. В случае биссектрисы $c_1 = cb/(a+b)$, $c_2 = ca/(a+b)$. Найдите сумму квадратов медиан.

9. Во всяком треугольнике $c = a\,\cos B + b\,\cos A$ (геометрически это очевидно, но можно вывести и из ОТП, сложив a^2 и b^2).

10. Во всяком треугольнике $\sin C = \sin A\,\cos B + \sin B\,\cos A$ — следует из 9 и теоремы синусов.

11. Тот факт, что сумма углов треугольника равна 180°, можно вывести из ОТП. (Это следует из утверждения задачи 10: $\sin A\,\cos B + \sin B\,\cos A = \sin(A+B)$, так что либо $\angle C = \angle A + \angle B$, либо $\angle A + \angle B = 180^\circ - \angle C$. Применяя тот же вывод к $\angle A$ и $\angle B$ вместо $\angle C$, убедимся, что $\angle A + \angle B + \angle C = 180^\circ$.)

Параллельные прямые. Завершим этот параграф традиционными теоремами о параллельных прямых. (Впрочем, как ясно из вышеизложенного, в элементарной планиметрии можно обойтись параллельными отрезками.) В заключение мы еще раз обсудим, более подробно, эквивалентные формулировки аксиомы параллельных.

Теорема 6. *Через каждую точку любой из двух прямых, перпендикулярных одному отрезку, проходит общий перпендикуляр этих прямых; все эти перпендикуляры равны друг другу.*

Это непосредственно следует из леммы 1, стоит лишь вместо отрезков с общим перпендикуляром представить себе содержащие их прямые. □

О прямых в теореме 6 говорят, что они расположены *на постоянном расстоянии* друг от друга (поскольку расстояние отмеряется по перпендикуляру). Прямые, которые не имеют общих точек, называются *параллельными.* Прямые, расположенные на постоянном расстоянии, очевидно параллельны. Верно также обратное; это вытекает из следующей «основной теоремы о параллельных».

Теорема 7. *Через каждую точку A, не лежащую на данном прямой a, проходит только одна прямая, параллельная a.*

Доказательство. Опустим из A на a перпендикуляр AB и проведем через A перпендикулярную ему прямую b. Она параллельна a.

Пусть b' — какая-либо проходящая через A прямая, отличная от b. Она пересекает b, и потому один из ее лучей c с началом A лежит с той же стороны от b, что и прямая a. Пусть α — угол, который луч c образует с перпендикуляром AB. Возьмем на c такую точку C, что $AB = AC \cos\alpha$. Это равенство означает, что AB — это проекция отрезка AC; стало быть, точка C лежит на прямой a, т. е. прямая b' ее пересекает, и теорема 7 доказана. □

Аксиома параллельных прямых. Обычно принимают не нашу аксиому о параллельных отрезках, а аксиому о параллельных прямых:

Через каждую точку, не лежащую на данной прямой a, проходит не более одной прямой, параллельной a.

А то, что есть хотя бы одна параллельная, доказывается (без всякой аксиомы о параллельных прямых или отрезка). Именно:

Через каждую точку A, не лежащую на данной прямой a, проходит прямая, параллельная a. Это прямая, перпендикулярная перпендикуляру AB к прямой a.

Выполняется и более общее утверждение, известное из курса средней школы (тоже доказываемое без всяких аксиом о параллельных):

Теорема 8. *Если прямые a, b, проходящие через точки A, B, образуют с отрезком AB с одной стороны от него углы, равные в сумме $2d$ — двум прямым углам, то эти прямые параллельны.*

Доказательство. Пусть прямые a, b, как сказано, образуют с отрезком AB углы с одной стороны, в сумме равные $2d$. Тогда с другой стороны — сумма углов та же самая (так как те углы смежные,

рис. 40). Допустим, вопреки утверждению, что прямые a, b пересекаются в какой-то точке C, так что имеем треугольник ABC.

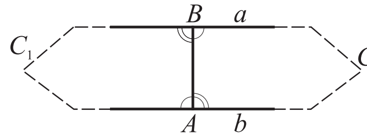


Рис. 40

Построим с другой стороны от AB треугольник ABC_1, равный $\triangle ABC$, но расположенный так, что его угол A равен углу B в $\triangle ABC$, и тогда его угол B равен углу A в $\triangle ABC$. Тогда при точках A и B углы обоих треугольников дают в сумме $2d$ — развернутый угол. Поэтому отрезки AC и AC_1 образуют один отрезок CC_1 так же, как отрезки BC и BC_1 образуют один отрезок CC_1.

Выходит, что точки C и C_1 соединены двумя отрезками. А это невозможно (по аксиоме). Следовательно, прямые a, b не пересекаются, что и требовалось доказать. □

То, что сумма односторонних углов A, B — с одной стороны от отрезка AB — равна развернутому углу, равносильно тому, что равны углы при A и B с разных сторон от AB — накрест лежащие углы (поскольку углы при A, как и при B, смежные и в сумме равны развернутому).

Поэтому из доказанного следует

Теорема 9. *Если накрест лежащие углы равны, то* $a||b$. □

Если принять аксиому параллельных, то можно доказать также обратное.

Теорема 10. *Если прямые a, b параллельны, то накрест лежащие углы при отрезке AB с концами на прямых a, b равны.*

Действительно, так как прямая b, проходящая через данную точку B и параллельная данной прямой a, только одна, то она и есть та, для которой накрест лежащий угол при точке B равен углу при A. □

Теорема 11. *Если принята аксиома о параллельных прямых, то выполняется аксиома параллельных отрезков: если отрезки AB и BD равны друг другу, лежат с одной стороны от отрезка AB и перпендикулярны ему, то $CD = AB$.*

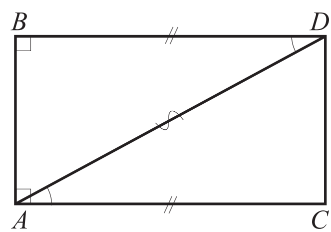


Рис. 41

Доказательство. Так как отрезки AC, BD перпендикулярны AB, то прямые AC, BD параллельны (по доказанному). Поэтому накрест лежащие углы при точках A и D равны (рис. 41). Следовательно, у треугольников ACD и BAD равны углы A и D и заключающие их стороны: AD общая и

$AC = BD$. Следовательно, $AB = CD$, что и требовалось доказать. $\square$

Таким образом, мы доказали, что при аксиоме параллельных прямых выполняется аксиома параллельных отрезков. А так как выше было доказано обратное, то аксиомы равносильны.

Постулат Евклида утверждает:

Если прямые a, b, проходя через концы отрезка AB, образуют с ним с одной стороны углы, в сумме меньше $2d$, то эти прямые пересекаются, причем с той же стороны от AB.

Оставим последнее уточнение — и без него постулат равносилен аксиоме параллельных. Действительно, все четыре угла, какие образуют прямые a, b с отрезком AB, составляют в сумме $4d$, так что либо с обеих сторон от AB суммы углов равны $2d$, либо с одной стороны сумма меньше $2d$. В первом случае прямые параллельны (по теореме 8), и, стало быть, во втором случае, если выполнена аксиома параллельных, они пересекаются, т. е. выполнен постулат Евклида.

Если же выполнен этот постулат, то во втором случае прямые пересекаются. Это значит, что прямая b, проходящая через точку B и параллельная a, только одна — та, для которой сумма углов с одной стороны как раз равна $2d$. Значит, выполнена аксиома параллельных.

Доказательство того, что при аксиоме параллельных выполняется постулат Евклида в полном объеме — с утверждением о том, с какой стороны пересекаются прямые, — мы оставляем читателю в качестве задачи. Понятно, задача состоит в возможно более непосредственном доказательстве, потому что если считать доказанной теорему о сумме углов треугольника, то утверждение, что прямые пересекаются с той стороны, где сумма углов меньше $2d$, становится очевидным. $\square$

§ 3. Начала стереометрии: прямые и плоскости в пространстве

Исходные факты, касающиеся прямых и плоскостей, изложены в § 8, гл. I. За ними следует рассмотрение взаимного расположения двух прямых, в частности, доказывается теорема: через точку вне данной прямой a проходит прямая, ей параллельная[8], и притом только одна.

Удобный способ дальнейшего изучения взаимного расположения

[8]Прямые в пространстве называются параллельными, если они лежат в одной плоскости и не имеют общих точек.

прямых и плоскостей дает применение векторов. Проследим это на основных примерах.

Векторы и операции с ними (скалярное произведение в том числе!) были определены в гл. III, ч. 1, отправляясь от направленных отрезков. Напомним:

Два направленных отрезка $\overrightarrow{AB}$, $\overrightarrow{CD}$ имеют одно направление — *сонаправлены*, если при продолжении они не расходятся, т. е. каковы бы ни были налегающие на AB и CD равные отрезки AH, CN, расстояние MN ограничено: меньше некоторой константы c.

Два направленных отрезка представляют один и тот же вектор (равны как векторы), если они сонаправлены и равны по длине.

Были доказаны:

Теорема 1. *Два направленных отрезка, сонаправленные с третьим, сонаправлены.*

Теорема 2. *Два направленных отрезка AB, CD сонаправлены тогда и только тогда, когда они либо лежат на параллельных прямых по одну сторону от прямой AC, либо лежат на одной прямой и один из лучей AB, CD содержит другой.*

Отсюда сразу выводится

Теорема 3. *Две прямые, параллельные третьей, параллельны.*

Доказательство. Пусть прямые a, b параллельны прямой c. Тогда, опираясь на теорему 2, можно, взяв на c направленный отрезок $\overrightarrow{CC_1}$, взять на a и b отрезки $\overrightarrow{AA_1}$, $\overrightarrow{BB_1}$, сонаправленные с ним. По теореме 1, они сонаправлены друг с другом. А стало быть, по теореме 2, они лежат на параллельных прямых, так что $a||b$, что и требовалось доказать. □

Точно так же, как в гл. III ч. 1, могут быть получены все результаты о векторах, которыми мы здесь будем пользоваться.

Далее на основе угла между векторами можно определить угол между прямой и плоскостью. В частности, вводится понятие перпендикулярности.

Перпендикуляры к плоскости.

Теорема 4. *Вектор, перпендикулярный к двум не коллинеарным векторам, лежащим в данной плоскости, перпендикулярен ко всякому вектору, лежащему в той же плоскости.*

Доказательство. Пусть $\boldsymbol{a}$, $\boldsymbol{b}$ — два неколлинеарных вектора в плоскости α и вектор $\boldsymbol{c}$ им перпендикулярен, так что $\boldsymbol{ac} = \boldsymbol{bc} = 0$.

Всякий вектор $\boldsymbol{p}$ в той же плоскости представляется как сумма

векторов $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, коллинеарных $\boldsymbol{a}$ и $\boldsymbol{b}$, так что $\boldsymbol{p}_1\boldsymbol{c} = \boldsymbol{p}_2\boldsymbol{c} = 0$. Поэтому

$$\boldsymbol{cp} = \boldsymbol{c}(\boldsymbol{p}_1 + \boldsymbol{p}_2) = \boldsymbol{cp}_1 + \boldsymbol{cp}_2 = 0,$$

т. е. $\boldsymbol{c} \perp \boldsymbol{p}$, что и требовалось доказать. □

Доказанная теорема очевидно равносильна следующей (в которой уже нет речи о векторах).

Теорема 5. *Прямая, перпендикулярная двум пересекающимся прямым, лежащим в данной плоскости, перпендикулярна всякой прямой, лежащей в этой плоскости, и тем самым* ***перпендикулярна*** *этой плоскости.* □

Этот признак перпендикулярности позволяет доказать, что:

(1) *Через любую данную точку проходит плоскость, перпендикулярная данной прямой, и притом только одна.*

(2) *Через всякую точку плоскости проходит прямая, перпендикулярная этой плоскости, и притом только одна.*

Читатель либо вспомнит, либо сам найдет доказательства этих утверждений.

(3) *Прямая, параллельная прямой, перпендикулярной плоскости, перпендикулярна той же плоскости.* Это можно вывести из того, что на параллельных прямых можно взять равные векторы и их скалярные произведения с векторами в плоскости равны.

(4) *Прямые, перпендикулярные одной плоскости, параллельны.* Выводится из (3).

Двугранный угол измеряется углом, образуемым лежащими на гранях отрезками, перпендикулярными ребру. Эти отрезки на каждой грани параллельны, и потому угол один и тот же, из какой точки ребра ни проведены отрезки (рис. 42).

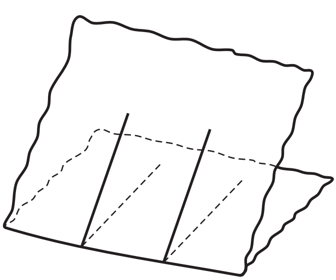


Рис. 42

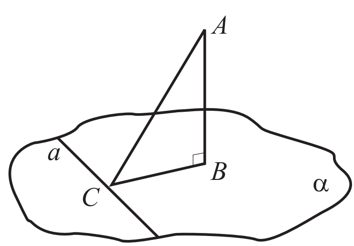


Рис. 43

Равенство углов между отрезками, представляющими те же самые векторы, лежит в основе определения скалярного произведения.

Три перпендикуляра.

Теорема 6. *Пусть A — точка вне плоскости α, и пусть прямая a и отрезок BC лежат на α, причем $C \in a$ и $BC \perp AB$* (рис. 43). *Возможны три перпендикулярности:*

а) $AB \perp \alpha$, б) $CA \perp a$, в) $CB \perp a$. *Утверждается, что если выполнены две из них, то выполняется и третья.*

Доказательство. Пусть $\boldsymbol{a} \neq \mathbf{0}$ — вектор на прямой a. Напишем скалярное произведение

$$\boldsymbol{a} \cdot \overrightarrow{AB} = \boldsymbol{a} \cdot (\overrightarrow{CB} - \overrightarrow{CA}) = \boldsymbol{a} \cdot \overrightarrow{CB} - \boldsymbol{a} \cdot \overrightarrow{CA}. \tag{1}$$

Поэтому если два из трех скалярных произведений $\boldsymbol{a} \cdot \overrightarrow{AB}$, $\boldsymbol{a} \cdot \overrightarrow{CB}$, $\boldsymbol{a} \cdot \overrightarrow{CA}$ равны нулю, то равно нулю и третье. А так как равенство нулю скалярного произведения равносильно взаимной перпендикулярности сомножителей, то получаем:

(а) Если $CA \perp a$ и $CB \perp a$, то $AB \perp a$. А так как по условию еще $AB \perp BC$ и $BC \subset \alpha$, то выходит, что вектор $\overrightarrow{AB}$ перпендикулярен двум векторам $\boldsymbol{a}$ и $\overrightarrow{BC}$, лежащим в плоскости α. Значит, $\overrightarrow{AB} \perp \alpha$.

(б) Если $AB \perp \alpha$ и $CA \perp a$, то $CB \perp a$.

(в) Если $AB \perp \alpha$ и $CB \perp a$, то $CA \perp a$. $\square$

Случай (а) дает способ опускания перпендикуляра из данной точки A на плоскость α, выраженный следующей теоремой. В этом заключено и доказательство существования опускаемого перпендикуляра.

Теорема 7. *Пусть A — точка вне плоскости α и a — прямая в α. Опускаем на a перпендикуляр AC и через C проводим в α прямую $b \perp a$. Перпендикуляр AB, опущенный на b, будет перпендикулярен к α.* $\square$

Случаи (б) и (в) дают следующую теорему (называемую теоремой о трех перпендикулярах).

Теорема 8. *Пусть из точки A на плоскость α опущен перпендикуляр AB, в плоскости α проведена прямая a и на ней взята точка $C(\neq B)$. Утверждается, что $AC \perp a$ тогда и только тогда, когда $BC \perp a$* (рис. 43). $\square$

Отрезок BC представляет собой проекцию отрезка AC — наклонной AC — на плоскость. Поэтому данная теорема кратко формулируется так:

Теорема 8′. *Прямая перпендикулярна наклонной тогда и только тогда, когда она перпендикулярна проекции этой наклонной.*

При $\boldsymbol{a} \cdot \overrightarrow{AB} = 0$ из формулы (1) следует, что

$$\boldsymbol{a} \cdot \overrightarrow{AC} = \boldsymbol{a} \cdot \overrightarrow{BC}. \tag{2}$$

Равенство (2) выражает даже больше: скалярное произведение единичного вектора $\boldsymbol{a}$ на какой-либо вектор $\boldsymbol{b}$ представляет собой проекции векторов $\overrightarrow{CA}$, $\overrightarrow{CB}$ на прямую a. Поэтому равенством (2) выражается

Теорема 9. *Пусть AC — наклонная к данной плоскости, BC — ее проекция на эту плоскость и a — прямая в этой плоскости, проходящая через точку C. Тогда проекции отрезков AC, BC на прямую a равны.* □

Параллельные плоскости — это плоскости, не имеющие общих точек. Их можно охарактеризовать так: *Плоскости параллельны тогда и только тогда, когда у них есть общий перпендикуляр.* (Докажите.)

Поэтому из существования перпендикуляра, опущенного на плоскость из данной точки, следует: *через каждую точку, не лежащую на данной плоскости, проходит плоскость, ей параллельная.*

Далее известна

Теорема 10. *Если плоскость α содержит две пересекающиеся прямые, параллельные двум прямым в плоскости β, то $\alpha || \beta$.*

Докажите ее разными способами, в частности, пользуясь перпендикулярами.

Из доказанных теорем можно вывести теорему, связанную с определением угла между прямой и плоскостью.

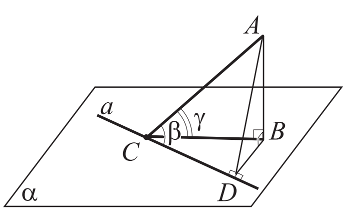


Рис. 44

Теорема 11. *Угол между наклонной к плоскости α и ее проекцией на α — наименьший из всех углов, образуемых наклонной с прямыми, проходящими в α через конец наклонной*[9].

Д о к а з а т е л ь с т в о. Пусть длина наклонной AC равна единице. Тогда длина ее проекции CD на прямую a, проходящую через C, равна косинусу угла β между AC и CD:

[9]То, что прямые проходят через конец наклонной, на самом деле не имеет значения, так как прямая, скрещивающаяся с данной, образует с ней, по определению, такой же угол, как прямая, ей параллельная.

$CD = \cos\beta$. А длина проекции BC наклонной на плоскость α равна косинусу угла γ между AC и BC (рис. 44). Итак,

$$\cos\beta = CD, \quad \cos\gamma = BC.$$

Вместе с тем, по теореме 9, CD есть проекция отрезка BC на прямую a. Поэтому $CD < BC$, и, следовательно, $\cos\beta < \cos\gamma$. А так как косинус — функция убывающая, то $\beta > \gamma$, что и требовалось доказать. □

§ 4. Фигуры с внутренними точками

Общие понятия границы и внутренности фигуры (множества точек) относятся к топологии (см. часть 5). Здесь мы определим их на плоскости и в пространстве, в понятиях нашей аксиоматики планиметрии и стереометрии. Роль, какую в планиметрии играет круг, в стереометрии играет шар.

Определение. Точка A считается *граничной* для данной фигуры, если сколь угодно близко к ней есть как точки, принадлежащие фигуре, так и точки не принадлежащие ей, т. е. те и другие точки есть во всяком круге (шаре) с центром A (рис. 45). Граничные точки фигуры образуют ее *границу*. Граничная точка может принадлежать фигуре, а может и не принадлежать ей. Множество (фигура), не содержащее ни одной своей граничной точки, называется *открытым*, а содержащее всю свою границу — *замкнутым*.

Точка фигуры, не являющейся ее граничной точкой, называется *внутренней*. Это, стало быть, такая точка, с центром в которой существует круг (шар), целиком содержащийся в фигуре (точки B, C на рис. 45). Внутренние точки фигуры образуют ее *внутренность*. О фигуре, содержащейся во внутренности фигуры F, говорят, что она содержится (лежит) *внутри* F.

Точку называют *внешней* для фигуры, если с центром в ней есть круг (шар), не содержащий точек фигуры (точка D на рис. 45). Это равносильно тому, что точка не принадлежит ни фигуре, ни ее границе.

Замечание. Подчеркнем, что на плоскости граничные и внутренние точки определяются с помощью кругов, описываемых вокруг точки, а в пространстве — с помощью шаров. Соответственно с этим граница и внутренность фигуры определяются на плоскости и в пространстве. Например, на плоскости круг имеет внутренние точки, и

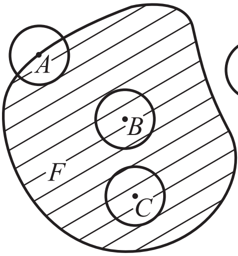


Рис. 45

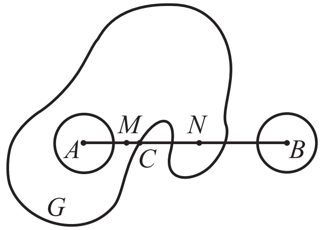


Рис. 46

его границей служит окружность. Но рассматриваемый как фигура в пространстве круг не имеет внутренности — все его точки граничные. Скажем, основание кругового цилиндра — это часть границы цилиндра. Вместе с тем мы говорим о точках *внутри* его основания, так же как о точках *внутри* граней многогранника, и т. п., имея в виду, что вокруг этих точек есть круги, содержащиеся в грани. (В аналогичном смысле говорят о точках «внутри» отрезка.)

Теорема 1. *Отрезок, соединяющий точку, лежащую внутри фигуры, с точкой, лежащей вне фигуры, пересекает границу фигуры* (*т. е. содержит хотя бы одну граничную точку*, рис. 46).

Доказательство. Пусть точка A лежит внутри фигуры G, а точка B — вне G. Проведем отрезок AB. На нем есть такие точки M, что отрезок AM содержится внутри G (поскольку вокруг A есть круг, содержащийся в G). Пусть a — супремум длин таких отрезков AM, и пусть AC — отрезок длины a. Точка C не лежит внутри G, так как иначе имелся бы содержащий ее отрезок внутри G и длина $a = |AC|$ не была бы супремумом длин отрезков AM, содержащихся внутри G. Точка C не лежит также вне G, так как иначе вокруг нее был бы круг (шар), не содержащий точек из G, и сколь угодно близко к C не было бы точек из G.

Таким образом, C не лежит ни внутри, ни вне G, а, стало быть, лежит на границе, что и требовалось доказать. □

Следствие. *Ломаная, соединяющая точку внутри фигуры с точкой вне фигуры, пересекает границу фигуры.*

Доказательство. Нужно воспользоваться индукцией по числу звеньев ломаной и применить предыдущую теорему 1. Подробности остаются читателю. □

Замечание. Доказанная теорема связывается с тем наглядным представлением, что граница фигуры отделяет ее внутренность от

внешних точек, или что она *ограничивает* фигуру.

Однако граница фигуры может очень мало соответствовать такому наглядному представлению, как показывают два следующих примера.

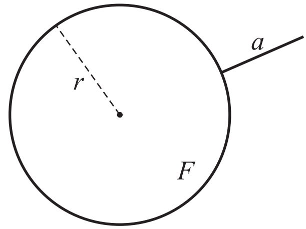


Рис. 47

Пример 1. Окружность круга служит его границей в наглядном смысле и ограничивает его. Но возьмем фигуру F, представляющую собой круг с исключенным радиусом r и с радиусом, продолженным на отрезок a (рис. 47). Граница этой фигуры представляет собой объединение окружности, ее радиуса r и отрезка a. В наглядном смысле трудно сказать, что такая граница ограничивает фигуру F.

Пример 2. Пусть F — множество точек с рациональными координатами x, y в квадрате Q: $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. Границей этого множества служит весь квадрат Q. Здесь нельзя сказать, что граница «ограничивает» фигуру F.

Замкнутая область. Выделим фигуры, границы которых не имеют таких особенностей, как в примерах 1, 2.

Определение. *Замкнутой областью* называется фигура со следующими свойствами:

(1) она обладает внутренними точками;

(2) она содержит свою границу;

(3) ее граница является также границей ее внутренности;

(4) любые две ее внутренние точки можно соединить ломаной, содержащейся целиком внутри фигуры.

Если фигура удовлетворяет условиям (1)–(3), то мы говорим, что ее граница *ограничивает* ее (и также ограничивает ее внутренность). Соответственно, мы говорим, что фигура F *ограничивает* фигуру G, если F служит границей фигуры G, удовлетворяющей условиям (1)–(3).

Например, окружность ограничивает круг, а сфера — шар; прямая ограничивает полуплоскость, плоскость ограничивает полупространство.

Определение. Фигура называется *ограниченной* (по размерам), если расстояния между ее точками ограничены, т. е. существует такой отрезок d, что всякий отрезок, соединяющий точки фигуры, по длине не превосходит d.

На плоскости это равносильно тому, что фигура содержится в некотором круге (достаточно взять круг радиуса d с центром в любой точке фигуры). В пространстве это равносильно тому, что фигура содержится в некотором шаре.

Согласно данным определениям, *замкнутая ограниченная область* — это ограниченная фигура с указанными выше свойствами 1–4.

В пространстве такая фигура называется ***телом*** (иногда условие ограниченности от тела не требуется). На плоскости такую фигуру назовем ***площадкой***.

Предостережение. Как это ни удивительно, одна и та же фигура может ограничивать не две, а три и... даже бесконечно много замкнутых областей. Соответственно, ограниченная фигура может на плоскости ограничивать несколько и даже бесконечное число площадок, а в пространстве — тел. К счастью, подобные «чудовища» водятся лишь в очень глубоких теоретико-множественных дебрях, и к «фигурам» их можно отнести лишь довольно условно. В случае многоугольных и многогранных фигур, к изучению которых мы переходим, ничего похожего не бывает. Тем не менее и здесь аккуратность в определениях не помешает.

Треугольник с внутренностью. Треугольник, понимаемый не как фигура из трех отрезков, а как ограниченная ими «часть плоскости», можно определить следующим образом.

Треугольник с внутренностью, или *плоский треугольник*, — это площадка, ограниченная тремя отрезками (т. е. граница ее представляет собой треугольник).

Теорема 2. *Плоским треугольником будет фигура, образованная отрезками, проведенными из вершины треугольника во все точки противоположной стороны. Из какой вершины проводятся отрезки — безразлично: фигура получается одна и та же.*

Доказательство. Внутренние точки указанной фигуры T — это те, которые лежат на отрезках, проведенных из вершины в точки на противоположной стороне. То, что вокруг каждой точки можно описать круг, содержащийся в T, очевидно, и доказательство мы оставляем читателям, так же как доказательство того, что границу фигуры T образуют стороны треугольника.

Нужно еще доказать, что фигура получается одна и та же, из какой бы вершины ни проводить отрезки до противоположной стороны. То есть если ABC — треугольник и M — точка на отрезке AD с концом D на BC, то M также лежит на отрезке BE с концом E на AC и на отрезке CF с концом F на AB (рис. 48).

Действительно, пусть точка D на BC и точка M на AD. Отрезки AD и AC служат поперечинами угла B. Поэтому отрезок BM при продолжении за точку M пересечет, согласно теореме 3 § 6, гл. I, сторону AC. Это завершает доказательство теоремы 2. □

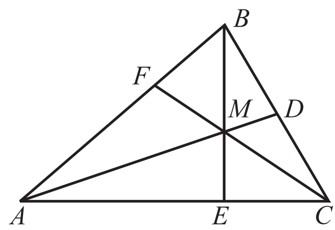


Рис. 48

Замечание. В определении плоского треугольника как площадки, ограниченной треугольником, не заключен способ определять, какие точки принадлежат площадке, а какие — нет. Доказанная теорема 2 дает такой способ отличать внутренние точки от внешних. Достаточно из какой-нибудь вершины провести через данную точку M отрезок, равный наибольшей из сторон треугольника. Он пересечет противоположную сторону тогда и только тогда, когда точка M внутренняя (докажите).

Отметим еще две теоремы.

Теорема 3. *Точка A лежит внутри треугольника тогда и только тогда, когда на стороне треугольника найдется такая точка B, что* луч AB *при продолжении за точку B не содержит точек сторон треугольника.*

Теорема 4. *Плоский треугольник является пересечением трех полуплоскостей, каждая из которых ограничена прямой, проходящей через две вершины, а сама содержит третью вершину.*

Доказательства оставляем читателю. □

Угол как «часть плоскости». *Плоским неразвернутым углом* называется замкнутая область (на плоскости), ограниченная двумя лучами, не образующими одну прямую. Эти лучи — *стороны* угла, их общее начало — *вершина* угла (поскольку лучи ограничивают область, то у них начало общее).

Плоский развернутый угол — это полуплоскость, у которой на ограничивающей ее прямой отмечена точка — вершина угла; лучи, на которые та делит прямую, — стороны угла.

Плоский угол обозначают его сторонами, если ясно, о каком из двух углов идет речь.

Теорема 5. *Два луча с общим началом ограничивают два плоских угла, при этом каждая точка, не лежащая на данных лучах, принадлежит одному и только одному из этих углов.* (Это очевидно... но, с точки зрения аксиоматического построения планиметрии, нуждается в доказательстве.)

Доказательство. Если два луча образуют одну прямую, то они ограничивают две полуплоскости. Поэтому для развернутых углов сказанное в теореме непосредственно следует из того, что прямая делит плоскость на две полуплоскости.

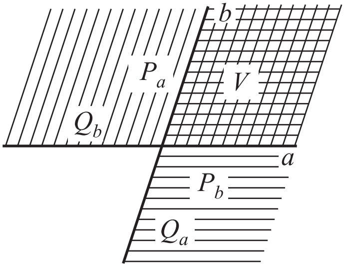


Рис. 49

Пусть теперь a, b — два луча с общим началом O, не образующие одной прямой, т. е. образующие неразвернутый угол ab. Пусть $\overline{a}$ и $\overline{b}$ — содержащие их прямые. Прямая $\overline{a}$ ограничивает две полуплоскости: одна из них содержит луч b — обозначим ее P_a, другую — Q_a. Аналогично, прямая $\overline{b}$ ограничивает полуплоскости P_b, Q_b, где $P_b \supset a$. Один плоский угол V, ограниченный лучами a, b, является пересечением полуплоскостей P_a, P_b, другой, W, является объединением полуплоскостей Q_a, Q_b (рис. 49).

Точка, принадлежащая лучу a, принадлежит прямой $\overline{a}$ и, стало быть, является граничной для полуплоскостей P_a, Q_a. Поэтому она граничная и для углов V, W. Аналогичное верно для точек луча b. Следовательно, лучи a, b служат общей границей углов V, W. Остальные точки угла V лежат внутри обеих полуплоскостей P_a, P_b, а остальные точки угла W — внутри хотя бы одной из полуплоскостей Q_a, Q_b. Следовательно, эти точки внутренние и для самих углов V, W. □

Таким образом, мы приходим к результату:

Теорема 6. *Два луча с общим началом ограничивают два плоских угла: один является пересечением, а другой — объединением полуплоскостей, ограниченных прямыми, содержащими эти лучи.*

(Формально это верно и для развернутого угла: в этом случае одна полуплоскость пересекается сама с собой, другая — объединяется сама с собой.) □

Определение. Плоский угол — объединение двух полуплоскостей — называется *сверхтупым.*

Замечание. Можно определить плоские неразвернутые углы несколько иначе. По отношению к двум лучам a и b, составляющим неразвернутый угол, все лучи, исходящие из вершины, кроме самих лучей a и b, делятся на два класса: класс F_1 тех лучей, которые проходят внутри угла ab, т. е. пересекает его поперечины, и класс F_2 всех остальных лучей. Плоские углы образуются лучами этих классов с

присоединением самих лучей a, b. (Это дает способ определять, какому из углов принадлежит данная точка.)

Определение. Плоский угол, вершина которого служит центром окружности, называется *центральным углом* этой окружности (рис. 50). Пересечение его с кругом называется *сектором* этого круга, и тогда угол называется *углом* этого *сектора*. Если угол развернутый, то сектор — полукруг.

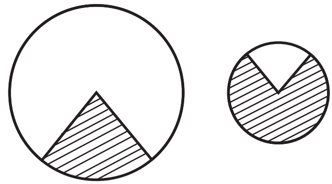


Рис. 50

Многоугольники и площадки. *Многоугольником* называют фигуру, образованную отрезками A_1A_2, A_2A_3, $\ldots$, A_nA_1 с условиями:

(а) эти отрезки не имеют общих точек, помимо концов, здесь они обозначены в круговом порядке (в частности, точки A_1, $\ldots$, A_n все различны);

(б) никакие два смежных отрезка не образуют одного отрезка. При этом условии отрезки называются *сторонами* многоугольника; их концы — его *вершины*. Поскольку вершин — n, то многоугольник называется *n-угольником*.

Многоугольником с внутренностью — короче, *плоским многоугольником* — называется площадка, границей которой служит многоугольник (ее *стороны* и *вершины* — это его стороны и вершины). Вообще, площадка, граница которой состоит из конечного числа отрезков, называется *многоугольной*. Но она может не быть плоским многоугольником, когда ограничена несколькими многоугольниками (рис. 51). Их вершины называются *вершинами* площадки.

Выполняется важная

Теорема 7. *Всякий многоугольник ограничивает плоский многоугольник (и при этом только один).*

Доказательство этой теоремы мы наметим в конце параграфа.

Всякий достаточно малый круг с центром в вершине плоского многоугольника делится сторонами, сходящимися в этой вершине, на два сектора, из которых только один содержится в этом многоугольнике (рис. 52). Угол этого сектора и называется *углом* многоугольника *при данной вершине*. У многоугольной площадки несколько углов могут иметь общую вершину (рис. 51).

Впрочем, под углом плоского многоугольника (как и многоугольной площадки) понимают обычно не столько фигуру, сколько соответствующую величину.

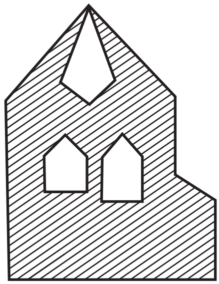


Рис. 51

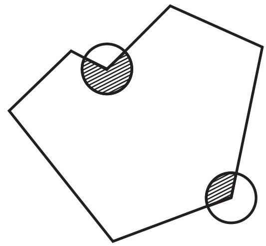


Рис. 52

Замечание. Термин «многоугольник» употребляют в разных смыслах, называя им (1) многоугольник как он определен выше; (2) вообще замкнутую ломаную, не исключая пересечений; (3) многоугольник с внутренностью; (4) многоугольную площадку, тогда многоугольник в нашем смысле называют *простым многоугольником.*

Многогранные углы. Многогранники. *Многогранным углом* называется фигура, образованная конечным числом плоских углов с общей вершиной и попарно общими сторонами, которые все можно обойти от одного к другому — смежному по стороне, — причем эти углы не имеют других общих точек, кроме как принадлежащих общим сторонам (рис. 53). Если обозначить плоские углы их сторонами, то можно сказать, что многогранный угол — это фигура, образуемая плоскими углами a_1a_2, a_2a_3, $\ldots$, a_na_1, причем эти углы не имеют общих точек, помимо принадлежащих их общим сторонам a_1, a_2, $\ldots$, a_n (которые попарно различны). Это определение аналогично определению замкнутой ломаной или многоугольника.

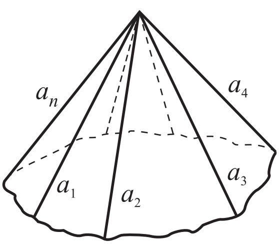


Рис. 53

Если смежные по стороне углы лежат в одной плоскости, то их можно объединить, так что соседних углов, лежащих в одной плоскости, не останется. Если никакие два соседних угла не лежат в одной плоскости, то углы называются *гранями* многогранного угла. Углы, образующие многогранный угол, могут быть развернутыми и сверхтупыми. (В частности, двугранный угол с точкой, отмеченной на его ребре в качестве вершины, подпадает под данное определение

многогранного угла. Но его не принято числить среди многогранных углов.)

Определение. *Многогранник* (как поверхность) — это фигура, образованная конечным числом многоугольных площадок так, что выполнены два условия:

(а) Если точка A является вершиной одной из площадок, то плоские углы всех площадок, которым эта точка принадлежит, образуют вокруг нее многогранный угол; при этом если точка A лежит на стороне какой-либо площадки, то она считается вершиной развернутого угла в этой площадке.

(б) От каждой площадки к любой другой можно перейти по смежным площадкам, т. е. имеющим общую сторону или отрезок стороны.

Смежные площадки, лежащие в одной плоскости, можно заменить на одну площадку. Это, очевидно, не нарушит условий (1), (2).

Площадка, не лежащая в одной плоскости ни с какой смежной, называется *гранью* многогранника. Точка, являющаяся вершиной какой-либо грани, называется *вершиной* многогранника. Сторона или отрезок стороны площадки, соединяющий две вершины (но не содержащий вершины), называется *ребром.* У многогранника на рис. 54 грань P — не многоугольник, и у многогранного угла в вершине A один плоский угол больше развернутого.

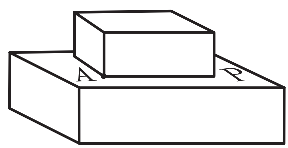


Рис. 54

Из условия (а), определяющего многогранник, следует, что для взаимного расположения двух граней имеются только следующие возможности: (1) они не имеют общих точек, (2) имеют одну общую вершину, (3) имеют одну общую сторону, (4) имеют несколько общих вершин и лежат в одной плоскости, (5) имеют несколько общих вершин и, может быть, несколько общих отрезков сторон, причем все эти вершины и отрезки содержатся в одной прямой (рис. 55).

Вообще, многогранник может иметь сложное строение — гораздо более сложное, чем привыкают представлять себе, имея в качестве примеров только пирамиды, призмы, простейшие выпуклые многогранники.

Многогранник называется *выпуклым*, если он располагается по одну сторону от плоскости каждой своей грани (не считая, конечно, самой грани). Грани выпуклого многогранника — выпуклые многоугольники. Каждые две либо не имеют общих точек, либо имеют общую

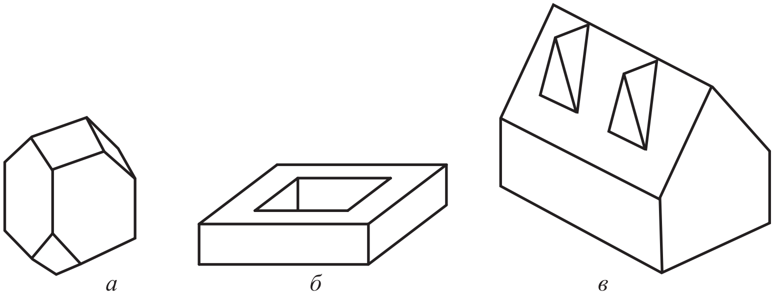


Рис. 55

сторону — ребро,— либо только одну общую вершину. Но на рис. 55, *б*, *в* даны примеры многогранников совсем другого строения.

Многогранником с внутренностью, или *телесным многогранником*, называют тело, ограниченное многогранником в смысле данного здесь определения. И выполняется

Теорема 8. *Всякий многогранник ограничивает телесный многогранник (и притом только один!).*

Доказательство наметим в конце параграфа.

Замечание. Многогранником (или полиэдром) называют также вообще любое тело, ограниченное конечным числом многоугольных площадок. (Такое тело уже может не быть телесным многогранником!) Для различения можно говорить «многогранное тело». Такие тела могут, вообще, иметь очень сложное строение — полости и другие особенности; мы их не перечисляем, а иллюстрируем на рис. 56. Ребром тела называется сторона или отрезок стороны ограничивающей его площадки, по которому она смежна с площадкой, лежащей в другой плоскости. Смежные площадки, лежащие в одной плоскости, образуют одну *грань*, если только они не разделены таким ребром — одним или несколькими (площадки P, Q на рис. 56 разделены ребром).

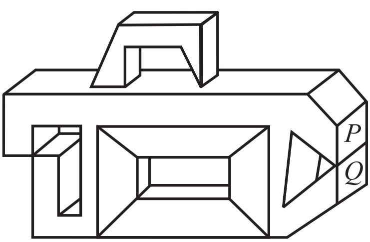


Рис. 56

Многогранниками называют также вообще многогранные поверхности, составленные из конечного числа многоугольных площадок, обычно с условием, что у каждой площадки любая данная ее сторона

либо общая у нее с одной и только одной площадкой, либо «свободна» — не принадлежит никакой другой площадке. Кроме того, требуется, чтобы все площадки с любой данной вершиной можно было обойти одну за другой, переходя только через их стороны, содержащие эту вершину. Такой многогранник представляет собой поверхность с краем, образованным свободными сторонами площадок. При таком понимании термина «многогранник» многогранник в смысле первого данного выше определения называют *замкнутым.*

О различении внутренних и внешних точек. Как уже было замечено в связи с определением плоского треугольника, само определение внутренних и внешних точек не заключает в себе способа их различать. Пока нет способа, позволяющего установить различие за конечное число операций, оговоренных в определении фигуры, мы не имеем основания признать плоский многоугольник фигурой в элементарной геометрии.

Мы укажем такой способ, затем наметим доказательство теоремы 7 о том, что всякий многоугольник ограничивает плоский многоугольник, и доказательство аналогичной теоремы для многогранников. При этом мы уже не будем следовать аксиоматической строгости, а положимся на очевидность.

Следующая теорема дает способ различать между собой внутренние и внешние точки плоского многоугольника.

Теорема 9. *Точка является внутренней для данного плоского многоугольника тогда и только тогда, когда любой проведенный из нее луч, не заключающий никакой стороны ограничивающего многоугольника, пересекает его в нечетном числе точек.* (При этом считается, что луч, проходя через вершину, пересекает многоугольник только в том случае, если сходящиеся в вершине стороны лежат по разные стороны от него.) *В частности, плоский многоугольник определен своей границей однозначно.*

Доказательство наглядно очевидно. Луч, исходящий из внутренней точки, должен выйти из многоугольника, пересекая его границу в какой-то точке. Далее он может опять войти в многоугольник, но тогда он должен из него выйти. Поэтому общее число пересечений вместе с первым будет нечетным.

Если же луч исходит из внешней точки, то, войдя в многоугольник, он должен из него и выйти. Поэтому число пересечений будет четным.

(Для проверки нет надобности проводить бесконечный луч, достаточно провести отрезок длиной не меньше полупериметра многоугольника.) □

Доказательство теоремы 7. Пусть дан многоугольник P. Выберем какой-нибудь луч a, не параллельный никакой стороне многоугольника; лучи, сонаправленные с a, назовем *допустимыми.*

Все точки плоскости, не принадлежащие самому многоугольнику, разделим на два класса G и H. В класс G отнесем такие точки M, что допустимый луч, идущий из M, пересекает многоугольник P в нечетном числе точек. В класс H отнесем все другие точки, т. е. такие, что исходящий из них допустимый луч пересекает многоугольник P в четном числе точек.

Точки класса G в совокупности с точками многоугольника P и образуют плоский многоугольник, ограниченный многоугольником P. Это и надо показать. Прежде всего нетрудно убедиться, что многоугольник P является границей фигуры, образованной точками класса G — в частности, она есть открытое множество. Остается доказать, что эта фигура ограничена и что любые две ее точки соединены в ней ломаной. Доказательство довольно просто, и мы оставляем его читателю. $\square$

Обратимся к многогранникам.

Теорема 10. *Точка является внутренней для телесного многогранника тогда и только тогда, когда любой проведенный из нее луч, не задевающий ребер и не проходящий через вершины ограничивающего многогранника, пересекает его грани в нечетном числе точек. В частности, телесный многогранник однозначно определяется своей границей.*

Доказательство совершенно аналогично доказательству теоремы 9. $\square$

Доказательство теоремы 8. Оно проводится по тому же плану, что и доказательство теоремы 7. Правда, возникающие при этом технические трудности более серьезны. Восстановление деталей и подробный вывод мы оставляем читателю в качестве (не очень простого, но интересного и поучительного) упражнения. $\square$

Заключительные замечания. Теоремы 9 и 10 практически дословно переносятся на случай многоугольных площадок и многогранных тел соответственно. Можно сформулировать и доказать подобные аналоги и для теорем 7, 8. Но здесь еще нужно точно описать граничные фигуры — т. е. границы площадок и тел. (Кстати, сделать это проще, чем дать определения многоугольника и многогранника.) Все это мы тоже предоставляем дотошному читателю.

Отметим еще, что всякая многоугольная площадка и, в частности, всякий плоский многоугольник составляются из треугольников, точно

так же, как всякое многогранное тело и, в частности, всякий телесный многогранник составляется из тетраэдров. (Точный смысл этих утверждений наглядно ясен.)

§ 5. Отображения. Наложения; их общие свойства

Отображения. В геометрии рассматривают разнообразные отображения фигур в пространстве и на плоскости, в частности отображения всего пространства или плоскости.

Отображение фигуры F в фигуру G состоит в том, что каждой точке фигуры F ставится в соответствие точка — одна единственная — из фигуры G. О точке M', соответствующей точке M при отображении f, говорят, что она является *образом* точки M, и пишут $M' = f(M)$. Если все точки фигуры G являются образами точек фигуры F, то говорят, что фигура F отображается *на* G.

Отображение по определению сопоставляет точке единственную точку, но может сопоставлять одну и ту же точку разным точкам. Так происходит, например, при проектировании на плоскость, когда одной точке N проекции фигуры F соответствуют все точки этой фигуры, лежащие на проектирующей прямой, проходящей через N. Если при данном отображении такого нет и разные точки отображаются на разные точки, то отображение называется *взаимно однозначным.*

Пусть отображение f фигуры F на G взаимно однозначно. Тогда каждая точка N из G соответствует одной (единственной) точке M из F. И обратно: каждой точке N из G соответствует одна точка M из F. А это значит, что задано отображение фигуры G на F, *обратное* отображению f. Его обозначают f^{-1}. Отображение, имеющее обратное, называют *обратимым.*

Из данных определений непосредственно следует, что если отображение f обратимо и f^{-1} — обратное ему отображение, то само f — обратное для f^{-1} (потому что если f сопоставляет точке M точку N, то f^{-1} сопоставляет точке N точку M).

Пусть заданы два отображения: отображение f фигуры F на G и отображение g фигуры G на H. Если при отображении f точка M из F отобразилась в точку N из G, а затем при отображении g точка N отобразилась в точку P из H, то тем самым M отобразилась на P. Это записывается так: $P = gf(M)$ (сначала f отображает M в N, а потом g отображает N в P). В результате получается отображение фигуры F на H. Это отображение h называют *композицией* отображения f с последующим отображением g и пишут $h = gf$ или $h = g \circ f$. Ком-

позицией называют как само отображение h, так и последовательное применение нескольких отображений.

Тривиальным примером отображения является *тождественное отображение* фигуры самой на себя — такое, когда каждой точке фигуры сопоставляется она сама. Если произвести сначала отображение f фигуры F на G, а потом обратное отображение f^{-1}, то каждая точка фигуры F отображается обратно в саму себя. То есть получим тождественное отображение. Обозначая его id_F, можно записать $f^{-1}f = \mathrm{id}_F$; и совершенно так же $ff^{-1} = \mathrm{id}_G$.

При рассмотрении отображений вместо того, чтобы говорить «точка M отображается в точку N», сплошь и рядом говорят «точка M переходит в N». Хотя она никуда не переходит. Нельзя двигать точки пространства; тем не менее наглядно отображения нередко представляют как перемещение и деформацию фигуры, как результат перемещения ее точек.

Неподвижной точкой отображения называют такую точку, которую оно «оставляет на месте», т. е. отображает саму в себя: $f(A) = A$.

Пример. Поворот круга вокруг центра представляет его взаимно однозначное отображение на себя, и центр является неподвижной точкой этого отображения. Обратным отображением будет поворот в обратную сторону на тот же угол.

Взаимно однозначное отображение фигуры на себя называют ее *преобразованием.* Впрочем, преобразованиями называют и отображения на другую фигуру; говорят, например, что окружность «преобразуется» в эллипс при сжатии.

Замечание. Отображение фигуры (множества) на фигуру (на множество) называется *сюръективным* отображением или, коротко, *сюръекцией.* Взаимно однозначное отображение называется *инъективным* отображением или, короче, — *инъекцией.* Взаимно однозначное отображение одного множества на другое называют *биективным*, короче, — *биекцией.*

Наложения. *Наложением* фигуры F мы называем[10] ее отображение, при котором каждым двум точкам A, B сопоставляются такие точки A', B', что отрезки $A'B'$, AB равны: $A'B' = AB$.

Точки, на которые отображаются точки фигуры F, образуют некоторую фигуру F'. И мы говорим: фигура F *налагается на* F'.

Тождественное отображение фигуры F очевидно является ее наложением самой на себя.

[10] Употребляются также термины «движение» и «перемещение».

Наложение, очевидно, взаимно однозначно: разные точки отображаются на разные, поскольку $AB = A'B'$. Поэтому наложение обратимо. Для него есть обратное отображение.

Отображение, обратное наложению, — тоже наложение. (Потому что если при прямом отображении точкам A, B отвечают A', B', то при обратном — точкам A', B' отвечают точки A, B, а так как $A'B' = AB$, то так же $AB = A'B'$.)

Композиция наложений является наложением.

Действительно, пусть при наложении f фигуры F на F' точки A, B отображаются в A', B', а при наложении фигуры F' на F'' точки A', B' переходят в A'', B''. Тем самым точки A, B отображаются в A'', B''. А так как при наложениях $A'B' = AB$ и $A''B'' = A'B'$, то $A''B'' = AB$, т. е. получается наложение фигуры F на F''.

Свойства наложений. Наложение по определению сопоставляет парам точек A, B такие пары A', B', что отрезки AB и $A'B'$ равны. Оно поэтому сохраняет все геометрические соотношения. Перечислим ряд основных свойств наложения. В последующих формулировках о наложении явно не говорится: оно подразумевается.

1. *Наложение сохраняет расстояние между точками, измеренное в каком бы то ни было масштабе.* Это служит только иным выражением того, что точкам A, B сопоставляются такие точки A', B', что $A'B' = AB$.

2. *Если точки A, B, C отображаются* в A', B', C', то $\angle ABC = \angle A'B'C'$ (как следует из определения равенства углов).

3. *Если точки A, B, M, N отображаются в A', B', M', N' и M принадлежит отрезку AB, то $M' \in A'B'$, а если $N \notin AB$ то $N' \notin A'B'$.*

Доказательство. Это следует из того, что точка C принадлежит отрезку $|AB|$ тогда и только тогда, когда $|AC| + |CB| = |AB|$. Подробнее: если $M \in AB$, то

$$|AM| + |MB| = |AB|.$$

А тогда по сохранению расстояний

$$|A'M'| + |M'B'| = |A'B'|,$$

и, значит, $M' \in AB$.

Если $N \notin AB$, то $|AN| + |NB| \neq |AB|$, и тогда также $|A'N'| + |N'B'| \neq |A'B'|$, т. е. $N' \notin A'B'$.

4. *Образом отрезка является отрезок.*

Доказательство. Пусть концы отрезка AB отобразились в точки A', B'. Тогда по предыдущему любая точка $X \in AB$ отображается в точку $X' \in A'B'$. И если $Y' \in A'B'$, то ввиду равенства $AB = A'B'$ на отрезке AB есть такая точка Y, что $AY = A'Y'$. Это точка отобразится на Y'. Таким образом, весь отрезок $A'B'$ оказывается образом AB.

5. *Образом плоского треугольника служит плоский треугольник.* Это следует из 4, поскольку треугольники заполняются отрезками, соединяющими вершину с точками противоположной стороны.

6. *Образом луча служит луч и образом прямой — прямая.* Это следует из 4, так как луч, как и прямая, представляется как объединение отрезков.

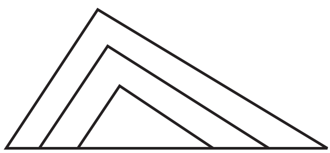


Рис. 57

7. *Образом плоскости служит плоскость, а полуплоскости — полуплоскость.*

Действительно, плоскость можно представить как объединение расширяющихся треугольников, а полуплоскость — как объединение расширяющихся треугольников с основаниями на прямой, ограничивающей полуплоскость. Поэтому сказанное следует из 5 и 6 (рис. 57).

8. *Образы параллельных прямых и плоскостей параллельны.* Для прямых ссылаемся на п. 6.

9. *Образом тетраэдра служит тетраэдр, образом пространства — пространство, образом полупространства — полупространство.*

Доказывается аналогично 5 и 7.

§ 6. Равенство фигур

Фигуры называются *равными* (или *конгруэнтными*), если между их точками существует такое взаимно однозначное соответствие, что отрезки, соединяющие соответственные точки, равны: если точкам A, B соответствуют A', B', то $AB = A'B'$.

То же определяют через наложение одной фигуры на другую. Иначе говоря:

Фигура F равна F', если существует наложение фигуры F на F', короче: если можно F наложить на F'.

Согласно определению наложения, это значит, что каждой паре точек A, B фигуры F соответствует такая пара A', B' в фигуре F', что

$A'B' = AB$. Ввиду обратимости наложения каждой паре точек C', D' из F' соответствует такая пара точек C, D из F, что $CD = C'D'$. Соответствие пар точек, стало быть, взаимно. И, следовательно, как фигура F' равна F, так и обратно: F равна F'. Таким образом, фигуры F и F' равны друг другу. Поэтому второе определение равенства фигур через наложение равносильно первому.

Замечание. Однако логически эти определения существенно различны. В первом обе фигуры играют одинаковую роль: они равны *друг другу*; отношение их равенства симметрично. Во втором фигуры играют разную роль: в нем одна фигура F отображается — налагается — на другую F', и тогда говорится, что F' равна F. Но то, что верно также обратное: F равна F', доказывается. Другими словами, симметричность равенства не выражена в определении, а доказывается.

Ввиду равносильности обоих определений равенства фигур можно дальше пользоваться любым из них. И из свойств наложений следует:

Две фигуры, равные третьей, равны друг другу.

Каждая фигура, очевидно, равна сама себе, поскольку каждой ее точке A можно сопоставить ту же A.

В итоге нами получена

Теорема 1. *Отношение равенства фигур рефлексивно, симметрично и транзитивно, т. е. является отношением эквивалентности.*
□

Равенство отрезков, углов и т. д. До данного здесь общего определения равенства фигур мы пользовались понятиями равенства отрезков, равенства углов и равенства треугольников. Нужно выяснить отношение этих понятий равенства к общему определению.

Начнем с равенства отрезков. Оно понимается у нас двояко. Во-первых, равенство отрезков введено в аксиоматике как основное отношение, оно тем самым определяется аксиомами, и по понятию не является равенством отрезков как фигур.

Во-вторых, поскольку было доказано, что всякий отрезок является фигурой, то к отрезкам применимо понятие равенства фигур: можно говорить о равенстве отрезков в смысле равенства их как фигур; когда между их точками устанавливается должное соответствие (при аксиоматически понимаемом равенстве отрезков ни о каком соответствии их точек нет речи).

Однако оба понятия равенства оказываются равносильными.

Теорема 2. *Два отрезка равны как фигуры тогда и только тогда, когда они равны в аксиоматическом смысле.*

Доказательство. Пусть отрезки AB, CD равны аксиоматически.

Между точками отрезка AB и числами x из промежутка $[0,\ |AM|]$ имеется взаимно однозначное соответствие, сохраняющее порядок и расстояния (см. теорему I. 4.4).

Если отрезок CD равен AB, т. е. $|CD| = |AB|$, то для него есть такое же соответствие. Поэтому если мы сопоставим точкам $M \in AB$ такие точки $N \in CD$, что $|AM| = |CN|$, то и получим соответствие, при котором отрезкам MM_1 соответствуют равные им отрезки NN_1.

Следовательно, равные отрезки AB, CD равны как фигуры.

Докажем обратное утверждение.

Пусть отрезки AB и CD равны как фигуры. Тогда, в частности, паре A, B отвечает такая пара A', B', что $A'B' = AB$. Если бы отрезок $A'B'$ не совпадал с CD, то он был бы меньше CD. А тогда для C, D не нашлось бы в отрезке AB такой пары C_1, D_1, что было бы $C_1D_1 = CD$! Отрезки AB, CD не могли бы быть равными как фигуры. Следовательно, паре концов A, B необходимо соответствует пара C, D, т. е. $AB = CD$, что и требовалось доказать. $\square$

Поскольку равенство фигур можно определять наложением, то доказанную теорему можно сформулировать так:

Отрезки накладываются один на другой тогда и только тогда, когда они равны (в аксиоматическом смысле).

Равенство углов. Угол, как он был определен, не является фигурой. Но можно с ним сопоставить фигуру, фиксировав на его сторонах определенные отрезки, отложенные от вершины. Если у двух углов

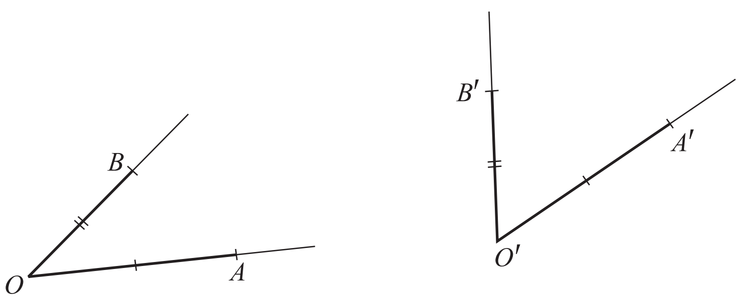


Рис. 58

отложить на сторонах соответственно равные отрезки, то их можно сравнивать как фигуры (рис. 58). Это приводит к следующей теореме.

Теорема 3. *Пусть на сторонах двух углов O, O' отложены равные отрезки $OA = O'A'$, $OB = O'B'$, так что получаются две фигуры — объединения этих отрезков по два. Углы O, O' равны тогда и только тогда, когда равны эти фигуры.*

Доказательство. Равенство углов определено равенством соответствующих поперечин, и у равных углов все соответственные поперечины равны. А если вспомнить, какие поперечины называются соответственными, то становится очевидным, что равенство углов равносильно равенству фигур, составленных из равных отрезков на их сторонах. □

Теорема 4. *Равенство треугольников, определенное по равенству соответственных сторон и углов, равносильно их равенству как фигур.*

Доказательство. По теореме 2 равенство сторон в аксиоматическом смысле равносильно их равенству как фигур, а по теореме 3 равенство углов равносильно равенству соответствующих фигур. □

Теорему 3 можно еще формулировать как теорему о наложимости.

Теорема 3а. *Углы равны тогда и только тогда, когда, один накладывается на другой «с точностью до укорочения или удлинения сторон», т. е., иначе говоря, если взять их стороны равными.* □

Аналогично выполняется теорема о треугольниках.

Далее, пользуясь доказанными теоремами и определением плоского треугольника (с внутренностью), нетрудно доказать, что плоские треугольники, ограниченные равными треугольниками, равны; так сказать, соответствие пар точек $MN = M'N'$, устанавливаемое на границе, распространяется внутрь.

§ 7. Площадь и ее применения

Обоснование измерения площади представляется еще более сложным, чем обоснование измерения длины отрезков. Поэтому мы также откладываем это обоснование до раздела «Основания геометрии», а здесь примем существование площади как аксиому. То, что у каждого многоугольника есть определенная площадь, принимается в школьном изложении без всяких оговорок, как нечто само собой разумеющееся. Но если подходить более серьезно, то это нужно формулировать явно, в виде аксиомы, если мы оставляем это без доказательства, как существование длины. Для формулировки аксиомы введем понятие многоугольной фигуры на плоскости.

Определение. *Многоугольной фигурой* назовем объединение конечного числа многоугольных площадок. Будем говорить, что фигура F *составлена* из фигур $F_1, \ldots, F_n$, если она служит их объединением и никакие две из них не имеют общих внутренних точек.

Аксиома площади. *Если некоторому квадрату E отнесено число единица, то каждой многоугольной фигуре F можно сопоставить положительное число — ее* ***численную площадь*** $S(F)$ ***в масштабе*** E *— так, чтобы были выполнены условия:*

а) *равные фигуры имеют одинаковую площадь;*

б) *если фигура F составлена из фигур F_1, F_2, то ее площадь равна сумме их площадей.*

Свойство б) называется *аддитивностью.* (Здесь, как и дальше, говоря «площадь», мы подразумеваем ее численное значение в каком-либо масштабе.)

Обратим внимание на полную аналогию этой аксиомы с аксиомой длины II_2 (§ 4 гл. I). Второе условие на длину в ней можно сформулировать так:

б′) *Если отрезок a составлен из отрезков a_1, a_2, то его длина равна сумме их длин.*

Аналогично можно пересказать условие в аксиоме III_3 о мере угла (§ 6 гл. I).

За квадрат E, служащий масштабом измерения площади, принимают квадрат со стороной, равной масштабу измерения длин, — «единичный квадрат»[11]. Тогда из аксиомы площади легко выводится:

Теорема 1. *Площадь прямоугольника равна произведению длин его сторон.*

Доказательство. Прямоугольник независимо от его расположения задается длинами сторон. И так как у равных фигур площади равны, то площадь прямоугольника P является функцией длин его сторон a, b: $S(P) = f(a,\ b)$.

Фиксируем какое-либо b и рассмотрим прямоугольники, у которых одна сторона равна этому b. Для них определим величину

$$g(a) = \frac{f(a,\ b)}{f(1,\ b)}. \tag{1}$$

[11]В принципе это не обязательно, но существенно тем, что позволяет выражать площади через длины (как, например, площадь прямоугольника через длины сторон) без множителей, зависящих от масштаба. В практике единицы площади согласуются с единицами длин не так просто; например, 1 гектар — это площадь квадрата со стороной 100 метров.

Если прямоугольники со сторонами a_1, b и a_2, b «прижать» друг к другу стороной, равной b, то получится прямоугольник со сторонами a_1+a_2, b. По аддитивности площади $f(a_1+a_2,\ b) = f(a_1,\ b)+f(a_2,\ b)$. Поэтому

$$g(a_1 + a_2) = g(a_1) + g(a_2).$$

Кроме того, очевидно, $g(1) = 1$, так что функция g удовлетворяет всем условиям, определяющим длину отрезка, и, стало быть (по теореме I.4.1 единственности длины), $g(a) = a$. Поэтому из (1)

$$f(a,\ b) = f(1,\ b)a. \tag{2}$$

Стороны a, b играют одинаковую роль. Поэтому совершенно так же $f(a,\ b) = f(a,\ 1)\cdot b$, откуда

$$f(1,\ b) = f(1,\ 1)b. \tag{3}$$

Но $f(1,\ 1)$ — это площадь квадрата со стороной единица, так что $f(1,\ 1) = 1$. Поэтому из (3) $f(1,\ b) = b$, и из (2) получаем

$$f(a,\ b) = ab,$$

т. е. $S(P) = ab$, что и требовалось доказать. □

Теперь мы можем доказать единственность численной площади и теорему о замене масштаба — такие же, как в случае длины и меры угла. Ограничимся единственностью (тем более, что из нее нетрудно вывести формулу замены масштаба). Поскольку каждая многоугольная фигура составляется из треугольников, то достаточно доказать, что численная площадь треугольника зависит только от масштаба, т. е. вычислить ее. Это мы сейчас и сделаем.

Теорема 2. *Площадь треугольника равна половине произведения основания на высоту.*

(Эта краткая условная формулировка означает: численная площадь треугольника равна половине произведения длины какой угодно его стороны на длину проведенной к этой стороне высоты (при условии, что площадь измеряется в масштабе единичного квадрата).)

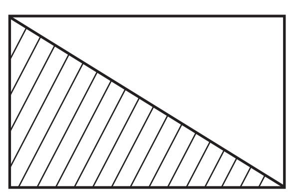


Рис. 59

Доказательство. Прямоугольный треугольник (рис. 59) представляет собою

«половину» прямоугольника, поэтому его площадь равна половине произведения его катетов.

Пусть теперь дан какой-либо треугольник ABC. Проведем из вершины C высоту CH — перпендикуляр к прямой AB. Возможны три случая:

1) точка H совпала с одной из вершин A или B (рис. 60, *а*);

2) точка H лежит на стороне AB (рис. 60, *б*);

3) H лежит на продолжении стороны AB так, что, например, B на AH (рис. 60, *в*); если A на BH, то можно изменить обозначение A на B.

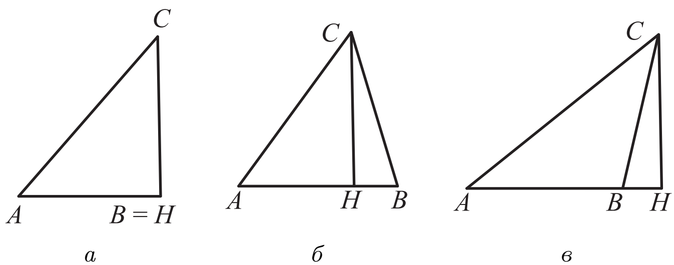


Рис. 60

В первом случае треугольник прямоугольный, и его площадь $S = \frac{1}{2}|AB| \cdot |CH|$.

Во втором случае треугольник слагается из двух прямоугольных, и

$$S = \frac{1}{2}|AH| \cdot |CH| + \frac{1}{2}|BH| \cdot |CH| = \frac{1}{2}|AB| \cdot |CH|.$$

В третьем случае $\triangle ABC$ получается из $\triangle AHC$ «вычитанием» $\triangle BHC$, и потому

$$S = \frac{1}{2}|AH| \cdot |CH| - \frac{1}{2}|BH| \cdot |CH| = \frac{1}{2}|AB| \cdot |CH|. \quad \square$$

Следствие. *Произведение стороны на опущенную к ней высоту одно и то же для всех трех сторон треугольника:* $ah_a = bh_b = ch_c$ (так как оно равно удвоенной его площади). $\square$

Приложения. Выражение для площади треугольника позволяет легко получить выражение для площадей параллелограмма, трапеции, правильного многоугольника — вообще для любой многоугольной площадки, поскольку ее можно разбить на треугольники.

Но главное — это возможность еще раз получить ряд основных результатов планиметрии, не имеющих прямого отношения к площади (ср. § 2), — прежде всего доказать теорему Пифагора и ввести тригонометрические функции.

Теорема Пифагора. *Для всякого прямоугольного треугольника площадь квадрата, построенного на гипотенузе, равна сумме площадей квадратов, построенных на катетах.*

Доказательство непосредственно получается из сопоставления рисунков 61, *а* и 61, *б*. (То, что на рисунке 61, *б* у четырехугольника, вписанного в квадрат со стороной $a + b$, углы прямые, следует из того, что каждый из них получается вычитанием из развернутого суммы острых углов прямоугольного треугольника.) □

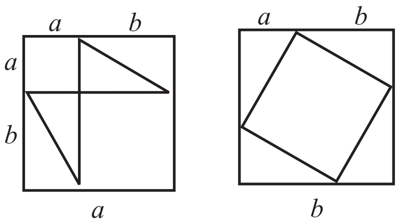


Рис. 61

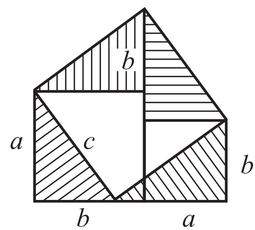


Рис. 62

Другое доказательство видно на рис. 62. (Квадрат, построенный на гипотенузе, составляется из тех же частей, что и фигура из квадратов, построенных на катетах.) □

Синус и косинус. *Синус* угла в прямоугольном треугольнике есть отношение противолежащего катета к гипотенузе. Но для того чтобы такое определение было правомерным, должна быть доказана

Теорема 3. *В прямоугольном треугольнике отношение катета к гипотенузе зависит только от величины противолежащего катету угла.*

Доказательство. Пусть в треугольниках ABC, $A_1B_1C_1$ углы C, C_1 — прямые и углы A, A_1 равны (рис. 63). На катетах $AC = b$, $A_1C_1 = b_1$ отложим равные отрезки $AD = A_1D_1 = d$. Проведем перпендикуляр $DE = h$ к гипотенузе $AB = c$. Если точка E_1 — такая точка на A_1B_1, что $A_1E_1 = AE$, то треугольники ADE и $A_1D_1E_1$ равны (по двум сторонам и углу между ними). Поэтому отрезок D_1E_1 перпендикулярен A_1B_1.

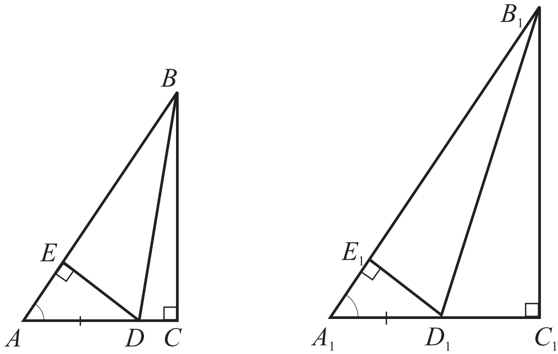


Рис. 63

Рассмотрим треугольник ABD. У него DE — высота, опущенная на AB, а $BC = a$ — высота, опущенная к AD. Поэтому

$$|AB| \cdot |DE| = |AD| \cdot |BC|, \quad \text{т. е.} \quad ch = ad,$$

откуда

$$\frac{a}{c} = \frac{h}{d}. \tag{4}$$

Рассмотрев треугольник $A_1B_1D_1$, совершенно так же получим, что (так как $|A_1D_1| = d$, $|D_1E_1| = h$)

$$\frac{a_1}{c_1} = \frac{h}{d}.$$

Вместе с (4) это дает

$$\frac{a_1}{c_1} = \frac{a}{c},$$

что и требовалось доказать. $\square$

Доказанная теорема 3 означает, что отношение противолежащего катета к гипотенузе есть функция величины угла. Это и есть синус угла:

$$\frac{a}{c} = \sin A. \tag{5}$$

Катет b, прилежащий к углу A, противолежит углу B, и поэтому $b/c = \sin B$. А так как $\angle B = 90^\circ - \angle A$, то $\sin B$ является также функцией

величины угла A. Эта функция — отношение прилежащего катета к гипотенузе — есть *косинус* угла A:

$$\frac{b}{c} = \cos A. \tag{6}$$

Рассматривая наклонную c, ее проекцию b и опущенный из ее конца перпендикуляр a, можно в определениях (5), (6) синуса и косинуса иметь в виду эти величины. И затем определить синус и косинус для тупых углов, а также для нулевого, прямого и развернутого угла, придавая при этом проекции знак минус, когда угол больше прямого (рис. 64).

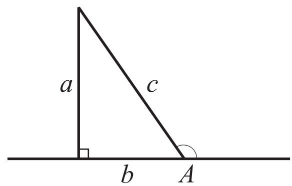


Рис. 64

В само́й элементарной геометрии нужны только углы, не большие развернутого. На определении синуса и косинуса для сверхтупых углов и далее — для отрицательных и больших 360° — мы сейчас не останавливаемся.

§ 8. Площадь и объем

В предыдущем параграфе было дано определение площади для многоугольных фигур, но в элементарной геометрии определяют площадь и других фигур: круга, кругового сектора и др. Трудность состоит в том, что нельзя приписать площадь со свойствами, указанными в «аксиоме площади», любым фигурам, нужно выделить фигуры, для которых это возможно. Это можно сделать, например, так.

Определение. Назовем площадку *простой*, если она ограничена конечным числом «криволинейных отрезков» — таких кривых, каждая из которых взаимно однозначно проектируется на прямолинейный отрезок (как сами прямолинейные отрезки, дуги окружности, части синусоиды и др., см. рис. 65).

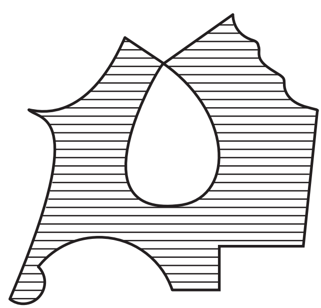


Рис. 65

В аксиоме площади можно заменить многоугольные фигуры на «простые» фигуры, т. е. составленные

из конечного числа простых площадок. Так получаем «обобщенную аксиому площади».

По этой аксиоме, если простая фигура F составлена из двух F_1 , F_2, то ее площадь $S(F) = S(F_1)+S(F_2)$, и, значит, $S(F) > S(F_1)$, поскольку $S(F_2) > 0$. Это обосновывает вывод площади круга из площадей вписанных и описанных многоугольников[12]. Как и выше, единственность площади доказывается. Аналогично можно определить объем.

Определение. Назовем тело *простым*, если оно ограничено конечным числом *криволинейных площадок*, т. е. таких фигур, каждая из которых взаимно однозначно и взаимно непрерывно проектируется на простую плоскую площадку. В частности, кривая площадка сама может быть простой плоской площадкой. Круговые цилиндры и конусы, шары, шаровые сегменты дают примеры «простых тел».

Аксиома объема. *Если некоторому кубу E отнесено число единица, то каждой фигуре F, составленной из конечного числа простых тел, можно сопоставить положительное число $V(F)$ —* ***численный объем фигуры*** F ***в масштабе*** E, *— и притом так, что выполнены условия:*

а) *равные фигуры имеют один и тот же объем;*

б) *если фигура F составлена из фигур F_1, F_2,*[13] *то ее объем равен сумме их объемов* (*аддитивность*);

За куб E примем «единичный куб», т. е. куб с длиной ребра, равной единице в принятом масштабе длин. Площадь будем измерять единичными квадратами.

Теорема 1. *Объем простого прямого цилиндра, основанием которого является простая фигура, равен произведению площади основания на высоту*[14].

Доказательство. Цилиндр определяется своим основанием и длиной образующей — для прямого цилиндра это то же, что высота.

---

[12] Площадь можно вообще определить таким способом: фигура F имеет площадь $S(F)$, если $S(F)$ есть точная верхняя и, одновременно, точная нижняя граница площадей многоугольных фигур, заключающихся в F и содержащих F. Но тогда надо доказывать, что так определенная площадь обладает свойствами аддитивности. Для круга это было бы не нужно. Но как без этого вывести площадь сектора? (Ведь сектор можно составить из других секторов.)

[13] То есть F есть объединение F_1 и F_2, и F_1, F_2 не имеют общих внутренних точек (в пространстве).

[14] *Цилиндром* здесь называется фигура, образованная равными параллельными отрезками (*образующими*), концы которых заполняют какую-либо плоскую фигуру (*основание*), причем образующие лежат с одной стороны от плоскости. Цилиндр *прямой*, если образующие перпендикулярны плоскости основания.

Поэтому объем $V(C)$ прямого цилиндра C зависит только от основания B и высоты h в данном масштабе:

$$V(C) = V(B, h).$$

Рассмотрим все прямые цилиндры с каким-либо заданным основанием B и образуем функцию

$$g(h) = \frac{V(B,\ h)}{V(B,\ 1)}. \tag{1}$$

Очевидно, $g(h) > 0$ и $g(1) = 1$. Два цилиндра с высотами h_1, h_2 составляются в цилиндр с высотой $h_1 + h_2$: его объем равен сумме их объемов. Поэтому

$$g(h_1 + h_2) = \frac{V(B,\ h_1 + h_2)}{V(B,\ 1)} = \frac{V(B,\ h_1)}{V(B,\ 1)} + \frac{V(B,\ h_2)}{V(B,\ 1)} = g(h_1) + g(h_2).$$

Таким образом, функция $g(h)$ удовлетворяет всем условиям, определяющим длину, и, стало быть, $g(h) = h$. Поэтому из (1)

$$V(B,\ h) = V(B,\ 1) \cdot h. \tag{2}$$

Убедимся, что величина $S(B) = V(B,\ 1)$ есть площадь основания B.

Действительно, очевидно, $S(B) > 0$, и если B есть единичный квадрат E_0, то прямой цилиндр с основанием E_0 и высотой $h = 1$ представляет собой единичный куб E, так что

$$S(E_0) = V(E_0,\ 1) = V(E) = 1.$$

Далее, если основания B, B' прямых цилиндров C, C' с равными высотами равны, то цилиндры равны. Поэтому если $B = B'$, то

$$V(B,\ 1) = V(B',\ 1), \quad \text{т. е.} \quad S(B) = S(B').$$

Наконец, если основание цилиндра C составляется из двух B_1, B_2, то цилиндр C составляется, соответственно, из двух цилиндров C_1, C_2 той же высоты. И его объем $V(C) = V(C_1) + V(C_2)$, так что

$$V(B_1 + B_2,\ 1) = V(B_1,\ 1) + V(B_2,\ 1),$$

т. е. $S(B_1 + B_2) = S(B_1) + S(B_2)$ (аддитивность).

Таким образом, величина $S(B) = V(B,\ 1)$ удовлетворяет всем условиям, определяющим площадь; она, стало быть, и есть площадь основания B, и из (2)

$$V(C) = V(B,\ h) = S(B) \cdot h,$$

что и требовалось доказать. □

Когда объемы прямых цилиндров определены, объем других тел можно находить, разбивая тело на тонкие слои, которые приближенно заменяются прямыми цилиндрами с простыми основаниями. Строгое оформление этой процедуры представляет интегрирование[15].

На этом пути доказываются и единственность объема, и формула замены масштаба. Интересно, что уже объем тетраэдра нельзя найти, не прибегая так или иначе к предельному переходу и не вводя дополнительных требований в аксиому объема — таких, например, как «принцип Кавальери».

Глава III

СПЕЦИАЛЬНЫЕ ВОПРОСЫ ЭЛЕМЕНТАРНОЙ ГЕОМЕТРИИ

§ 1. Задачи на построение

В задачах на построение речь идет о построении геометрических фигур данными средствами — обычно с помощью циркуля и линейки. При этом имеются в виду идеальные построения: пользование линейкой означает возможность проводить «прямые», т. е. отрезки произвольной длины; пользование циркулем означает возможность описывать окружность с любым центром любого радиуса. Эти возможности были выражены в «Началах» Евклида в качестве первых постулатов: 1) от одной точки до другой можно провести прямую; 2) всякую прямую можно продолжить; 3) вокруг каждой точки можно описать окружность любым радиусом.

Помимо построений циркулем и линейкой, можно, конечно, рассматривать построения с помощью других средств. Но такую возможность вместе с общим взглядом на геометрические построения мы обсудим в конце раздела о построениях, а сейчас сосредоточимся на построениях циркулем и линейкой. Простейший пример дает известное

[15] См., напр.: *Александров А. Д., Вернер А. Л., Рыжик В. И.* Геометрия: Пособие для школ с углубленным изучением математики. М.: Просвещение, 1987.

решение задачи: разделить отрезок пополам. Описываем вокруг концов данного отрезка AB окружности радиусами, равными AB, и соединяем их точки пересечения C, D. Существование этих точек, как и то, что и отрезок CD пересекает AB, строго говоря, нужно доказать. Это мы сейчас сделаем.

Теорема 1. *Если прямая содержит точку внутри круга, то она пересекает его окружность в двух точках.*

Доказательство. Пусть точка A прямой a лежит внутри круга с центром $O \notin a$ и радиусом R (рис. 66). Тем самым $|OA| < R$, и если OB — перпендикуляр к прямой a, то $|OB| \leqslant |OA|$, и, стало быть, $|OB| < R$. Если точки C, D на прямой a лежат с разных сторон от B и $|BC|^2 = |BD|^2 = R^2 - |OB|^2$, то $|OC| = |OD| = R$. Следовательно, в этих точках (и только в них!) прямая и пересекает окружность. $\square$

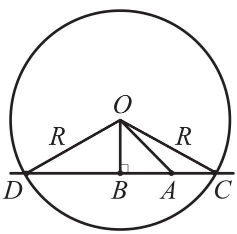


Рис. 66

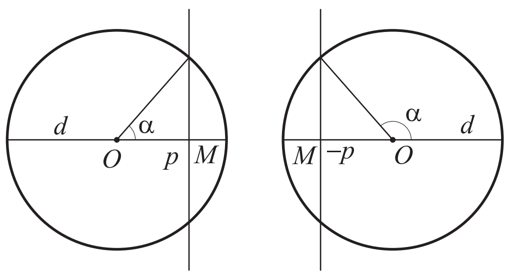


Рис. 67

При помощи этой теоремы может быть доказана

Лемма. *Косинус угла принимает все значения от* 1 *до* -1.

Воспользуемся известным определением косинуса с помощью окружности (рис. 67). Выберем единичный отрезок. Опишем окружность таким радиусом («единичную окружность»). Проведем в ней диаметр d. Косинус угла определяется как взятая с должным знаком длина проекции радиуса, образующего данный угол с «начальным радиусом». Расстояния OM от точек диаметра d до центра O принимают все значения от 0 до 1 (по теореме 4, § 4 гл. I). Поэтому если дано число p между 1 и -1, то берем точку M так, что $\pm|OM| = p$. Проводим через эту точку M прямую, перпендикулярную диаметру d. По теореме 1 она пересекает окружность. Проведя радиус в точку пересечения и получим угол α, у которого $\cos\alpha = p$. $\square$

Теперь рассмотрим пересечение окружностей.

Теорема 2. *Две окружности пересекаются (в двух точках) тогда и только тогда, когда расстояние между их центрами меньше суммы и больше разности их радиусов* (рис. 68).

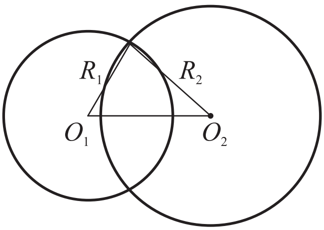


Рис. 68

Радиусы, проведенные в точку пересечения, вместе с отрезком, соединяющим центры, образуют треугольник. Поэтому высказанная теорема равносильна следующей.

Теорема 2а. *Три отрезка могут служить сторонами треугольника тогда и только тогда, когда какой-либо один из них меньше суммы и больше разности двух других.*

В треугольнике сказанное выполняется для каждой стороны в силу известной теоремы (см., например, теорему II.2.5). Поэтому вопрос состоит, собственно, в доказательстве обратной теоремы. Но мы докажем теорему 2а полностью.

Доказательство. В треугольнике со сторонами a, b, c и углом γ против c

$$c^2 = a^2 + b^2 - 2ab\cos\gamma, \quad \cos\gamma = \frac{a^2+b^2-c^2}{2ab}. \tag{1}$$

Откуда

$$1+\cos\gamma = \frac{a^2+b^2-c^2+2ab}{2ab} = \frac{(a+b)^2-c^2}{2ab}, \tag{2}$$

$$1-\cos\gamma = \frac{c^2-(a-b)^2}{2ab}; \tag{3}$$

и так как $-1 < \cos\gamma < 1$, то из (2) и из (3)

$$a+b > c > |a-b|. \tag{4}$$

Этим доказана известная теорема, что сторона треугольника меньше суммы и больше разности двух других. (Выведите (4) из (2) и (3).)

Докажем обратное утверждение. Пусть a, b, c — длины отрезков с неравенствами (4). Положим

$$\frac{a^2+b^2-c^2}{2ab} = p, \quad \text{т. е.} \quad c^2 = a^2+b^2-2abp.$$

Тогда подобно (2) и (3) получим

$$1 + p > 0, \quad 1 - p > 0, \quad \text{т. е.} \quad -1 < p < 1.$$

Так как косинус принимает все значения между 1 и —1, то существует такой угол γ, что $\cos\gamma = p$.

Треугольник с данными сторонами a, b и углом γ между ними существует (его можно построить). В таком треугольнике третья сторона как раз равна c. Теорема доказана. □

Основные построения. Перечислим основные задачи на построение, решения которых даются в школьном курсе (они будут дальше считаться известными без оговорки). Некоторые из этих построений мы показываем на рисунках с краткими указаниями в скобках.

1. *Отложить вдоль данного отрезка AB отрезок AE, равный данному CD.* (Описываем вокруг центра A окружность радиусом, равным CD, и, если нужно, продолжаем AB за точку B до пересечения с ней, рис. 69.)

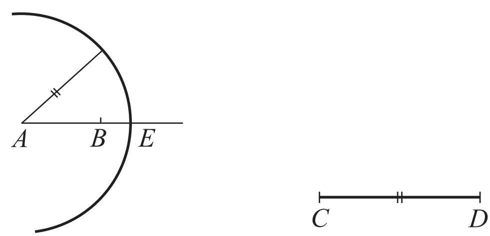


Рис. 69

2. *Построить сумму и разность данных отрезков.* (Даны AB и CD. Продолжая AB за точку B, откладываем на продолжении отрезок, равный CD. Получаем сумму. Разность — откладываем меньший на большем.)

3. *Отложить от данного отрезка угол, равный данному.* (Описав окружность вокруг вершины O данного угла, отсекаем на его сторонах равные отрезки OA, OB. Вокруг конца O_1 данного отрезка проводим окружность P того же радиуса, получаем отрезок $O_1A_1 = OA$, налегающий на данный. Вокруг A_1 описываем окружность Q радиусом AB. Если B_1 — точка пересечения P и Q, то O_1B_1 — сторона нужного угла.)

4. *Построить сумму и разность данных углов.* (Получается применением предыдущего построения.)

5. *Найти середину отрезка.* (Указано на с. 216 и 233.)

6. *Построить биссектрису угла.* (На сторонах угла откладываем от вершины равные отрезки OA, OB. Вокруг A и B проводим окружности, например, радиусом $|AB|$. Соединяем дальнюю точку пересечения и вершину.)

7. *Восстановить перпендикуляр к прямой в данной ее точке.* (Построение такое же, как в 6, потому что перпендикуляр — это биссектриса развернутого угла.)

8. *Опустить перпендикуляр из данной точки на данную прямую.* (Берем на данной прямой a точку B и вокруг данной точки A описываем окружность радиусом AB. Пусть она пересечет a в точке $C \neq B$. Дальше действуем так же, как при делении отрезка BC пополам. (А как быть, если окружность имеет с прямой a только одну общую точку B?))

9. *Провести через данную точку прямую, параллельную данной.* (На данной прямой a берем какую-нибудь точку B, соединяем ее с данной точкой A. Через A проводим прямую, образующую с AB такой же угол, как AB образует с данной прямой, но на противоположной стороне от AB (рис. 70). Другое построение, более короткое, можно основать на том, что параллелограмм — это четырехугольник с равными противоположными сторонами.)

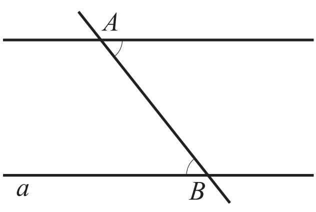


Рис. 70

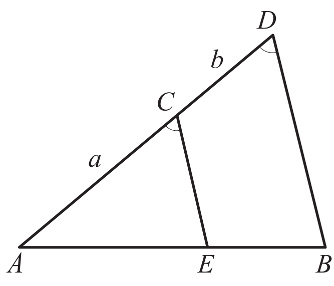


Рис. 71

10. *Разделить данный отрезок в данном отношении, заданном как отношение данных отрезков $a : b$ или как отношение целых чисел $n : m$.* (Построение ясно из рисунка: AB — данный отрезок, под углом отложим сумму отрезков $AC = a$ и $CD = b$. Если E на AB и $CE || DB$, то $AE : BE = a : b$. Если дано отношение $n : m$, то берем какой-либо отрезок c, строим $a = nc$, $b = mc$ и проводим изложенное построение (рис. 71). В частности, так получается построение данной доли $1/k$ отрезка AB, если взять $n = 1$, $m = k - 1$.)

Во всех рассмотренных построениях нужно (а) доказать, что они действительно дают нужный результат, (б) посмотреть, какие возможны варианты (например, при построении биссектрисы окружности вокруг A и B пересекаются в двух точках; что эти точки дадут?).

Решение задачи на построение. В нем различают четыре этапа:

1) поиск решения (анализ);

2) описание найденного решения или его фактическое выполнение (что, впрочем, обычно не требуется);

3) доказательство того, что найденное построение действительно решает поставленную задачу;

4) исследование решения.

В описании построения известные основные построения, как правило, просто указываются. Исследование состоит в выяснении того, всегда ли задача имеет решение; если не всегда, то при каких конкретных данных и сколько именно решений имеет задача.

При этом *разными* считаются решения, дающие неравные фигуры (или если и равные, то различно расположенные относительно данной фигуры, с которой связывается построение).

Задачи на построение треугольников.

1. *Построить треугольник по двум сторонам и углу между ними.* Решение очевидно; построение всегда возможно. Оно единственно в силу теоремы о равенстве треугольников.

2. *Построить треугольник по стороне и двум прилежащим углам.* Построение очевидно. Но оно дает результат лишь при условии, что сумма данных углов меньше развернутого. (Пятый постулат Евклида как раз и состоял в утверждении, что требуемое построение получится, если сумма данных углов меньше двух прямых.) Решение единственно (т. е. все такие треугольники равны друг другу).

2а. *Построить треугольник по стороне и двум углам, не обязательно прилежащим к данной стороне.* Третий угол находится путем вычитания двух данных из развернутого. Сколько возможно решений? Когда три, два, когда одно?

3. *Построить треугольник по трем сторонам.* Задача разрешима, если наибольший из данных отрезков меньше суммы двух других. Решение единственно.

4. *Построить треугольник по двум сторонам и углу, противолежащему одной из них.*

Решение. На одной из сторон данного угла откладывается отрезок, равный данному, представляющему собой сторону треугольника. Вокруг его конца описывается окружность радиусом, равным другой

стороне. Для этой окружности возможны 4 случая: а) она не имеет с другой стороной угла общих точек, б) они друг друга касаются, в) пересекаются в двух точках, г) пересекаются только в одной точке (рис. 72, *а, б, в, г*). Фиксация этих возможностей и точное указание того, при каких условиях реализуется каждая из них, и представляет

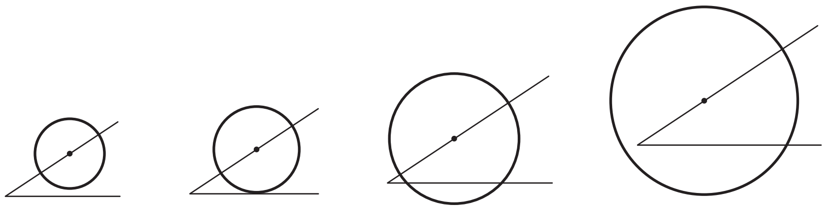


Рис. 72

исследование задачи. Проведите его.

5. *Построить треугольник по стороне, сумме двух других сторон и углу, прилежащему к данной стороне.*

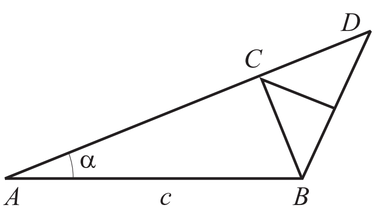


Рис. 73.

П о и с к р е ш е н и я (анализ). Допустим, задача решена, т. е. мы имеем треугольник ABC с данной стороной $AB = c$, суммой сторон $AC + CB = d$, углом $A = \alpha$ (рис. 73). Продолжив сторону AC на отрезок $CD = CB$, получим треугольник ABD со сторонами $AB = c$, $AD = d$ и углом $A = \alpha$. Вместе с ним получим также равнобедренный треугольник BCD; $BC = CD$.

Отсюда выводим решение задачи. Строим треугольник ABD с углами $A = \alpha$ и сторонами $AB = c$, $AD = d$. Находим на AD такую точку C, что $CD = CB$, т. е. что $\triangle BCD$ равнобедренный. Тогда треугольник ABC и дает решение задачи. Точку C можно найти, проводя перпендикуляр к отрезку BD через его середину или откладывая у точки B угол, равный $\angle D$ (пользуясь характерными свойствами равнобедренного треугольника: в первом случае совпадением медианы и высоты, во втором — равенством углов при основании).

И с с л е д о в а н и е. Сумма двух сторон треугольника больше тре-

тьей, поэтому должно быть $d > c$. При этом условии $AD > AB$, и потому $\angle D < \angle B$. Поэтому равнобедренный треугольник BCD можно построить, т.е. при построении угла DBC, равного $\angle D$, точка C будет на AD. При $d > c$ решение всегда возможно. Оно единственно, поскольку построение треугольника BCD проводится однозначно.

6. *Построить треугольник по стороне c, сумме двух сторон d и углу γ, противолежащему данной стороне.*

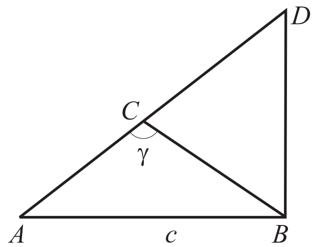


Рис. 74.

П о и с к р е ш е н и я. Допустим, имеется треугольник ABC с указанными данными: $AB = c$, $AC + CB = d$, $\angle C = \gamma$ (рис. 74). Продолжив отрезок AC на $CD = CB$, получим отрезок $AD = d$; затем построим треугольники ABD и BCD. Последний треугольник равнобедренный, и угол $C = \gamma$ для него внешний. Поэтому $\angle D = \frac{1}{2}\gamma$. Это приводит нас к следующему решению. Если построить треугольник ABD по сторонам $AB = c$, $AD = d$ и углу $D = \frac{1}{2}\gamma$, то, построив в нем равнобедренный треугольник BCD, найдем искомый треугольник ABC.

О п и с а н и е п о с т р о е н и я. Строим $\triangle ABD$ по данным сторонам $AB = c$, $AD = d$ и углу $D = \frac{1}{2}\gamma$. Это построение описано в решении задачи 4. Допустим, что оно проходит и мы получаем треугольник ABD. Затем строим равнобедренный треугольник BCD с вершиной C на AD (так же, как в предыдущей задаче). Получаем треугольник ABC, который и нужно было построить.

Д о к а з а т е л ь с т в о того, что он и есть искомый треугольник. Действительно, $AB = c$ и $AC + CB = AC + CD = d$, так как по построению $CB = CD$. Наконец, $\angle C = \gamma$, так как C — внешний угол равнобедренного треугольника BCD, в котором $\angle B = \angle D = \frac{1}{2}\gamma$.

И с с л е д о в а н и е. В задаче 4 треугольник получается не всегда, и, кроме того, могут получаться два треугольника. Что будет в условиях решаемой здесь задачи с построением треугольника ABD?

В условиях задачи должно быть $d > c$ — сумма двух сторон больше третьей. Поэтому в $\triangle ABD$ $AD > AB$. Кроме того, если $p = AE$ — перпендикуляр из A на прямую BD, то должно быть $AB \geqslant AE$, т.е. $c \geqslant p$. Данный угол $C = \gamma$ меньше развернутого, так что $\angle D = \frac{1}{2}\gamma$ — острый. Поэтому основание E перпендикуляра AE лежит на луче DB,

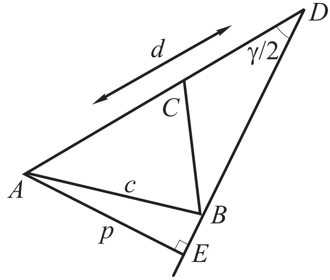


Рис. 75

а из прямоугольного треугольника AED имеем $AE = AD \sin D$, т. е. $p = d \sin \gamma/2$ (рис. 75).

Мы приходим к следующим случаям:

(а) Если $c < d \sin \gamma/2$, т. е. данная сторона c меньше катета прямоугольного треугольника с гипотенузой d и углом $\gamma/2$, то задача не имеет решения.

(б) Если $c = d \sin \gamma/2$, то получается один треугольник ADB — он же ADE — и, соответственно, один треугольник ACB.

(в) Если $d > c > d \sin \gamma/2$ — то получаются два треугольника ADB_1, ADB_2 (с $AB_1 = AB_2 = c$) и соответственно два треугольника AC_1B_1, AC_2B_2 (рис. 76).

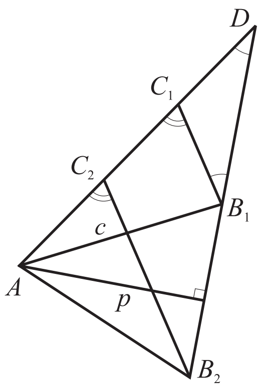


Рис. 76

Треугольники AC_1B_1, AC_2B_2 равны и отличаются только расположением сторон. В случае (б), когда $c = d \sin \frac{\gamma}{2}$, треугольник ACB один: он равнобедренный (почему?).

7. *Построить треугольник по двум углам и периметру.*

Р е ш е н и е. Треугольники с двумя данными углами подобны, их стороны пропорциональны и, следовательно, относятся как периметры. Поэтому, построив один такой треугольник, по нему получим подобный ему с данным периметром p.

Строим какой-нибудь $\triangle ABC$ с данными углами A, B. Откладываем вдоль луча AB отрезки AB, $BC' = BC$, $C'D = AC$. На луче AC откладываем отрезок AD_1, равный данному отрезку p, и известным приемом делим его на отрезки AB_1, B_1C_1, C_1D_1, пропорциональные AB, BC', $C'D$. Эти отрезки и дают стороны искомого треугольника (рис. 77).

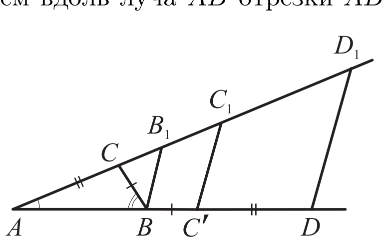


Рис. 77

Решение единственно, так как данная величина (отрезок) p един-

ственным образом разлагается на слагаемые, пропорциональные данным величинам (отрезкам).

Отметим, что при решении сколько-нибудь сложных задач на построение поиск решения задачи обычно состоит в следующем. Предполагаем, что требуемая фигура построена, и исследуем ее в связи с данными задачи с тем, чтобы найти путь от этих данных к тем, по которым фигура строится просто. Так и были решены задачи 5, 6, 7.

§ 2. Решение задач на построение

Метод подобия. Метод, примененный в решении последней задачи 7, — это «метод подобия». Он применим тогда, когда, исключив одно из условий задачи, можно построить фигуру, подобную искомой. Построив какую-нибудь такую фигуру, ее подобно преобразуют так, чтобы получить нужную фигуру.

Этот метод, в частности, всегда приложим, когда нужно построить треугольник по двум данным углам и еще по какой-нибудь линейной величине (по отрезку): ею может быть периметр (как в решенной задаче 7), сумма двух сторон, высота и т. д., и т. п.

Приведем пример другого рода задачи, решаемой методом подобия.

Задача. *Построить окружность, касающуюся сторон данного* (*неразвернутого*) *угла и проходящую через данную точку A внутри угла.*

Решение. Строим окружность, касающуюся сторон данного угла. Для этого откладываем на сторонах угла от его вершины O равные отрезки и в их концах проводим перпендикуляры к сторонам; в их пересечении получаем центр S окружности, он лежит на биссектрисе угла. Из вершины угла проводим луч через точку A. Он пересекает окружность в двух точках B, C. Находим на биссектрисе точки S_1, S_2 такие, что $OS : OS_1 = OB : OA$, $OS : OS_2 = OC : OA$. Это и будут центры окружностей с требуемыми свойствами.

Метод двух геометрических мест. Этот метод применим тогда, когда нужно построить точку M, удовлетворяющую двум условиям. Тогда строим, если возможно, два геометрических места точек, удовлетворяющих одному и другому из этих условий; в пересечении этих геометрических мест получаем точку M.

Применение этого метода можно видеть в простейших построениях, когда проводится окружность — геометрическое место точек, удаленных от данной точки на данное расстояние. Так, при построении треугольника по трем сторонам ищется его вершина C, находящаяся

на данных расстояниях a, b от двух других вершин — A, B. Описываются окружности вокруг A и B радиусами b и a, и в их пересечении получают вершину C.

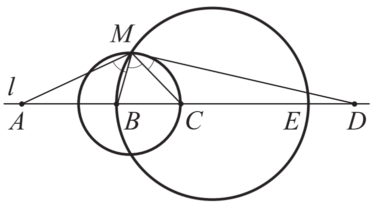


Рис. 78

Приведем интересный пример задачи, решаемой методом геометрических мест.

Задача. *На прямой l даны четыре точки A, B, C, D. Требуется найти такую точку M, из которой отрезки AB, BC, CD видны под равными углами, т. е. $\angle AMB = \angle BMC = \angle CMD$* (рис. 78).

П о и с к р е ш е н и я. Требование задачи можно перефразировать так, что в треугольниках AMC и BMD отрезки MB, MC должны быть биссектрисами. По известному свойству биссектрисы она делит противоположную сторону на отрезки, пропорциональные сторонам. Следовательно, должно быть

$$AM : CM = AB : CB, \quad BM : DM = BC : CD. \tag{1}$$

Первое равенство означает, что точка M принадлежит геометрическому месту точек, отношение расстояний которых от двух данных точек A, C равно $AB : CB$ (т. е. такое же, как у точки B). Это геометрическое место есть окружность (окружность Аполлония); § 3, гл. I, ч. 1) или прямая, если $AB = BC$. Совершенно так же точка M должна принадлежать геометрическому месту точек, отношение расстояний которых от точек B и D равно $BC : CD$. Это тоже окружность или прямая. Искомая точка M должна быть точкой пересечения этих окружностей или прямой с окружностью. (Прямые здесь не могут пересекаться, так как обе они перпендикулярны прямой, на которой лежат точки A, B, C, D. Таким образом, сразу видно, что задача не всегда имеет решение.)

П о с т р о е н и е. Положим $AB = a$, $BC = b$, $CD = c$. Окружность, описываемая точками M с первым условием (1), пересекает прямую l в точке B и еще в такой точке E, что $AE : CE = AB : CB = a : b$. Центр этой окружности — середина отрезка BE. Если $a > b$, то $AE = AB + BC + CE = a + b + x$ и

$$\frac{a+b+x}{x} = \frac{a}{b}, \quad x = b\,\frac{a+b}{a-b}.$$

Точка E лежит так, что C находится между B и E. Отсюда выводим построение, каким можно найти точку E и центр окружности. Если $b > a$, то точка E лежит с другой стороны, так что A находится между B и E.

Аналогичный вывод делаем для другой окружности.

Исследование. Задача интересна, в частности, условием ее разрешимости, которое выглядит несколько неожиданно.

Если средний отрезок b меньше крайних a и c, то решение есть (в крайности допустимо $a = b < c$ или $a > b = c$). Если b не меньше a и c, то решения нет. Оба вывода можно сделать из наглядных соображений (проведите их). Таким образом, если, скажем, $a < b$, то должно быть $c > b$. Вопрос в том, каковы точные соотношения между a, b и c.

Из расположения окружностей можно вывести, что они пересекаются, только если

$$3ac > b(a + b + c),$$

так что

$$a > b\frac{1 + b/c}{3 - b/c}; \quad c > b\frac{1 + b/a}{3 - b/a}.$$

Поэтому, в частности, решение вообще возможно только, если $a > b/3$, каким бы ни было при этом c. (Проведите детально все выводы.)

Метод преобразований. Этот метод состоит в том, что фигура, входящая в данные задачи, или ее часть подвергается преобразованию — переносу, симметрии и т. п., после которого решение получается просто. Приведем пример применения центральной симметрии.

Задача. *Дана точка A и две фигуры F_1, F_2. Требуется провести через точку A такой отрезок с концами в F_1 и F_2, чтобы точка A делила его пополам.* Разумеется, сами фигуры F_1, F_2 должны строиться циркулем и линейкой; в простейшем случае это прямая и окружность.

Решение. Строим фигуру F_3, симметричную F_2 относительно точки A. Если F_3 и F_1 имеют общую точку B, то отрезок BC с концом C, симметричным B относительно A, решает задачу. (Опишите построение фигуры F_3 и исследуйте задачу, когда фигуры F_1, F_2 — прямые, когда обе они — окружности, или одна — прямая, а другая — окружность.)

Следующая задача дает пример применения переноса.

Задача. *Даны две окружности K_1, K_2 и прямая a. Построить отрезок, равный данному d, параллельный a, с концами на K_1, K_2.*

Решение. Переносим окружность K_2 параллельно прямой a на отрезок d. Точка ее пересечения с K_1 и будет концом искомого отрезка (рис. 79). (Опишите построение в деталях. Между прочим, переносить можно в двух направлениях. Исследуйте задачу: выясните условия ее разрешимости, определите, когда сколько возможно решений.)

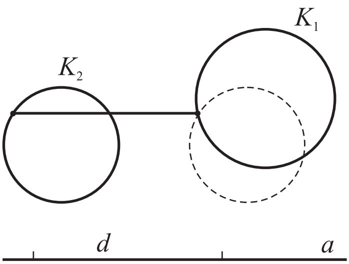


Рис. 79

Следующие две задачи на построение трапеций также решаются с помощью переноса. В первой из них переносится боковая сторона, во второй — диагональ. Вектор переноса задается меньшим из оснований трапеции.

Задача. *Построить трапецию по основаниям и боковым сторонам.*

Решение. Строим треугольник с основанием, равным разности данных оснований трапеции, и боковыми сторонами, равными боковым сторонам трапеции (рис. 80, a). Затем достраиваем его до трапеции.

Опишите построение и укажите условие разрешимости задачи.

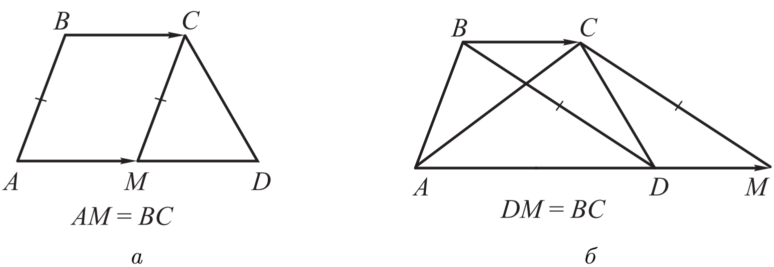


Рис. 80

Задача. *Построить трапецию по основаниям и диагоналям.*

Решение. Строим треугольник с основанием, равным сумме оснований трапеции, и с боковыми сторонами, равными данным ее диагоналям (рис. 80, $б$). Как из этого треугольника получить нужную трапецию, очевидно.

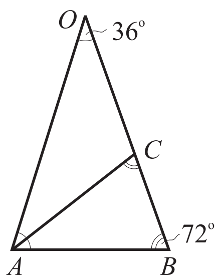
Исследуйте решение. Выясните, когда построение возможно. Убедитесь в единственности решения.

Правильные многоугольники. Задача построения правильного многоугольника обычно ставится так: *вписать в окружность правильный n-угольник* (с данным n).

Проще всего эта задача решается для шестиугольника, так как его сторона равна радиусу окружности. Треугольник вписывается в окружность по шестиугольнику — соединением его вершин через одну. Вписанный квадрат, очевидно, строится проведением двух взаимно перпендикулярных диаметров. Правильный пятиугольник строят по правильному десятиугольнику, соединяя вершины через одну.

Вписанный правильный десятиугольник строится так.

Допустим, построение выполнено. Сторона десятиугольника является основанием AB равнобедренного треугольника ABO с боковой стороной R и углом O при вершине в центре окружности O, равным $\frac{1}{10}360^\circ = 36^\circ$ (рис. 81, a). Поэтому углы при основании равны: $\frac{1}{2}(180^\circ - 36^\circ) = 72^\circ = 2 \cdot 36^\circ$.



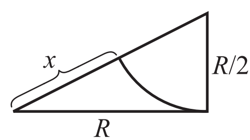


Рис. 81

Следовательно, биссектриса одного из этих углов A делит треугольник OAB на два треугольника AOC и ABC, являющихся равнобедренными (по равенству углов при основании). При этом у треугольника ABC угол B при основании такой же, как у $\triangle OAB$. Поэтому треугольники подобны и

$$\frac{BC}{AB} = \frac{AB}{AO}; \quad AB^2 = AO \cdot BC. \tag{2}$$

Кроме того, $BC = OB - OC = OA - AB$. Поэтому, полагая $AB = x$ (и так как $OA = R$), получаем из (2)

$$x^2 = R(R - x), \quad x^2 + Rx - R^2 = 0.$$

Положительный корень этого уравнения

$$x = \frac{-R + \sqrt{5R^2}}{2} = \frac{\sqrt{5} - 1}{2} R.$$

Отрезок длины x строится так. Строим прямоугольный треугольник с катетами R и $R/2$; его гипотенуза будет

$$\sqrt{R^2 + \frac{1}{4} R^2} = \frac{1}{2} \sqrt{5} R.$$

Вычитая отрезок $R/2$ получаем (рис. 81, *б*)

$$x = \frac{1}{2}(\sqrt{5} - 1) R.$$

Если построен правильный n-угольник, то правильный $2n$-угольник строится очевидным образом путем деления пополам центрального угла, опирающегося на сторону n-угольника. Так можно строить $2^m n$-угольник. В частности, по квадрату строится восьмиугольник, по шестиугольнику — двенадцатиугольник. Итак, циркулем и линейкой строятся правильные n-угольники с $n = 3,\ 4,\ 5,\ 6,\ 8,\ 10,\ \ldots$

Но для $n = 7,\ 9,\ 11,\ 13$ построение невозможно. Однако оно возможно для 17-угольника. Все это имеет алгебраическое основание, которое сейчас будет указано.

Вопрос о разрешимости задач на построение циркулем и линейкой. Коротко можно сказать так. Для того чтобы задача на построение циркулем и линейкой была разрешима, необходимо и достаточно, чтобы алгебраически она сводилась к решению уравнений, корни которых находятся посредством рациональных операций и операции извлечения квадратного корня (рациональные операции — это четыре действия арифметики: сложение, вычитание, умножение и деление).

Докажем необходимость. Пусть задача геометрически разрешима циркулем и линейкой. Введем на плоскости систему прямоугольных координат. То, что в задаче дано, и то, что надо построить, алгебраически задается координатами точек и длинами отрезков (как, скажем, координаты центра и радиус окружности). Решение задачи на построение сводится к нахождению точек пересечения (включая касание) прямых и окружностей, причем прямые задаются набором точек, через которые они проводятся, а окружности — своими центрами

и радиусами. Координаты точек пересечения находятся по этим данным посредством рациональных операций и извлечения квадратного корня, что и требовалось доказать (как фактически находятся координаты точек пересечения — полезно вспомнить).

Докажем обратное. Если задача на построение алгебраически решается с помощью рациональных операций и извлечения квадратного корня, то она разрешима циркулем и линейкой.

Для доказательства достаточно убедиться, что указанные алгебраические операции могут быть осуществлены посредством построения циркулем и линейкой. Эти операции должны производиться над длинами отрезков или, что равносильно, координатами точек. Геометрически им должны соответствовать построения с отрезками. Сложение и вычитание отрезков производится известным путем (вычитание большего отрезка из меньшего недопустимо в отличие от вычитания чисел, но это можно обойти откладыванием отрезков на разных полуосях координат).

Умножение и деление отрезков в прямом смысле невозможно. Но это обходят тем, что по трем данным отрезкам a, b, c можно построить отрезок $d = \frac{a}{c}b$, где отношение отрезков a/c — это то же, что отношение их длин. Отрезок d строится из подобия $d : b = a : c$ известным путем (рис. 82).

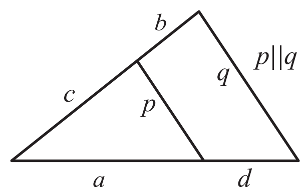


Рис. 82

Фиксируем отрезок e длины $|e| = 1$.

Если длины представлены отрезками a, b, то отрезок $d = \frac{a}{e}b$ представляет произведение длин $|a||b|$.

Аналогично, отрезок $d_2 = \frac{a}{b}e$ представляет a/b — частное длин $|a|$, $|b|$.

Пусть, наконец, нужно извлечь корень: найти $\sqrt{|a|}$. Число $|a|$ представлено отрезком a. Найдем такой отрезок d_3, что $|d_3|^2 = |a| \cdot |e|$. (Например, высоту в прямоугольном треугольнике с отрезками гипотенузы, равными a и e.) Тогда

$$|d_3| = \sqrt{|a||e|} = \sqrt{|a|}.$$

Таким образом, утверждение доказано. $\square$

Из доказанного становится достаточно очевидным, что разрешимость задачи на построение с помощью циркуля и линейки является редким исключением, поскольку разрешимость уравнений только с извлечением квадратного корня является исключением.

Построение правильного n-угольника алгебраически сводится к решению уравнения $x^n = 1$. Это уравнение имеет n корней, которые на

комплексной плоскости представляются точками на единичной окружности, служащими вершинами правильного n-угольника. Выражение корней в квадратных радикалах относится к вещественной и мнимой части. Как доказал Гаусс (в 1801 г.), оно возможно тогда и только тогда, когда

$$n = 2^m \cdot p \cdot \ldots \cdot p_s,$$

где m — любое неотрицательное целое число, а $p_1, \ldots, p_s$ — попарно различные простые числа, представимые в виде

$$p = 2^{2^k} + 1$$

с каким-либо целым k. При $k = 0,\ 1,\ 2,\ 3$ получаем $p = 3,\ 5,\ 17,\ 257$, поэтому 17-угольник строится циркулем и линейкой.

Три классические задачи рассматривались еще древними: трисекция угла, удвоение куба, квадратура круга. Все они не разрешимы циркулем и линейкой.

Первая задача состоит в том, чтобы разделить данный угол на три равные части. Эта задача приводится к кубическому уравнению и разрешима циркулем и линейкой лишь в исключительных случаях: например, прямой угол делится, но уже угол в 60° не делится на три равные части.

Задача удвоения куба состоит в том, чтобы построить по данному отрезку a такой отрезок x, чтобы куб с ребром x имел вдвое больший объем, чем куб с ребром a. Это сводится к уравнению $x^3 = 2a^3$, но $\sqrt[3]{2}$ в квадратных радикалах не выражается.

Задача квадратуры круга состоит в том, чтобы построить квадрат, площадь которого равна площади данного круга. Это очевидно приводится к уравнению $x^2 = \pi$. А так как π — число трансцендентное, то задача не разрешима никакими алгебраическими средствами.

Трансцендентность числа π доказал в 1882 г. немецкий математик К. Л. Ф. Линдеман (1852–1939).

О построениях вообще. Вообще под *построением* в геометрии понимают конечную последовательность действий, которые считаются возможными в силу оговариваемых условий; каждое из них дает какую-нибудь фигуру, — как отрезок, точку пересечения отрезков и т. п. При построениях циркулем и линейкой такими действиями являются проведение «прямых» и окружностей.

В чисто отвлеченном понимании, соответствующем аксиоматическому изложению, никаких «действий» быть не может. «Действие»,

дающее какую-то фигуру, обозначает только утверждение: «существует такая-то фигура». Так, мы говорим, например: «проведем отрезок AB», но это значит «существует отрезок с концами A, B». Построение, в отвлеченном понимании представляет собой последовательность таких утверждений, возможных в силу принятых условий. При построениях циркулем и линейкой эти условия утверждают существование прямых, проходящих через две данные точки, окружностей с любыми центрами и радиусами.

При аксиоматическом изложении геометрии условия, определяющие возможные построения, заключаются прежде всего в самих аксиомах. При принятой нами аксиоматике допустимыми построениями являются: проведение отрезка между двумя точками, откладывание отрезка, равного данному, и откладывание угла, равного данному. При этом, как уже сказано, эти «построения» не означают ничего, кроме утверждаемого аксиомами идеального, мыслимого существования объектов, о которых идет речь. Вместе с тем эти аксиомы выражают в отвлеченной форме совершенно реальные построения. Для них можно указать инструменты — это: 1) линейка, по которой можно проводить отрезки, отмечать на ней и переносить отрезки, равные данным, 2) угольник с подвижными сторонами и поперечиной для откладывания углов (рис. 83).

Можно доказать, что эти средства равносильны линейке и циркулю, т. е. всякая задача, решаемая посредством линейки и циркуля, решается также посредством линейки с возможностью отмечать отрезки и угольника, и обратно.

Выше была доказана теорема о существовании перпендикуляра к прямой (гл. II, § 1). Это доказательство представляет собой построение перпендикуляра, основанное на аксиомах — на проведении отрезков и откладывании угла. Вот еще один пример.

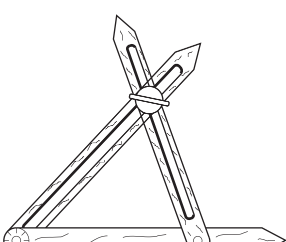


Рис. 83

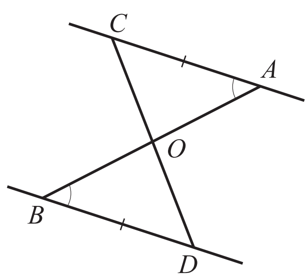


Рис. 84

Построение середины отрезка. Из конца A отрезка AB проводим отрезок AC под углом и из конца B — отрезок BD, равный AC и под равным углом к BA по другую сторону от AB (рис. 84). Тогда CD пересекает AB в середине. (Нужно прежде всего доказать, что CD пересечет отрезок AB, а не его продолжение. Дальнейшее сравнительно просто. Если воспользоваться тем, что уже известно существование середины O, то доказательство упрощается: проведите CO, DO.)

Заметим, что принятая нами аксиома существования отрезка данной длины дает основание откладывать при построениях отрезки данной длины. Это равносильно применению идеальной абсолютно точной измерительной линейки. Но это не соответствует, можно считать, характеру геометрических построений. Дальше, в части «Основания геометрии», будут даны аксиомы, в которые не входит понятие длины.

Построения в пространстве. Когда говорится о построениях в пространстве, то также имеют в виду не столько реальные построения, сколько утверждения о существовании объектов: «Проведем плоскость через точки A, B, C» означает: «фиксируем, что существует плоскость, проходящая через точки A, B, C». Аналогичный смысл имеет «проведение» прямых в пространстве.

Соответственно, построение фигуры фактически означает доказательство ее существования, основанное на аксиомах.

Рассмотрите основные построения в пространстве: 1) прямой, параллельной данной; 2) прямой, перпендикулярной данной (восстановление и опускание перпендикуляра); 3) прямой, перпендикулярной плоскости; 4) параллельных плоскостей и др.

§ 3. Выпуклые фигуры

Здесь мы будем считать понятия кривой и поверхности интуитивно ясными. Подробнее они обсуждаются в частях 4 и 5.

Выпуклые кривые. Будем рассматривать фигуры на плоскости. Ломаная называется *выпуклой*, если она расположена с одной стороны от каждой прямой, содержащей ее звено (рис. 85). Аналогично определяется *выпуклый* многоугольник как ломаная, а также как часть плоскости, ограниченная простой замкнутой ломаной (рис. 86).

Эти определения, известные еще из школьного курса, обобщаются на другие фигуры.

Определение. Прямая называется *опорной* прямой данной фигуры, если она имеет с фигурой или с ее границей хотя бы одну общую

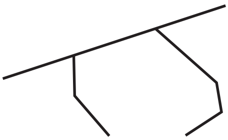


Рис. 85

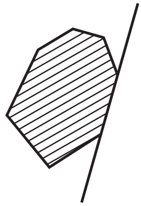


Рис. 86

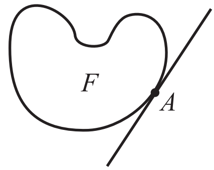


Рис. 87

точку, но вся фигура лежит с одной стороны от этой прямой, т. е. в одной ограниченной ею полуплоскости. Говорят: прямая *опорная* к фигуре F *в точке* A, если она при этом проходит через точку A фигуры F (рис. 87).

Название «опорная» выражает то наглядное представление, что фигура как бы опирается на прямую. Пользуясь понятием опорной прямой, можно пересказать данное выше определение выпуклой ломаной. Ломаная называется ***выпуклой***, если каждая прямая, содержащая ее звено, служит ее опорной прямой. Многоугольник ***выпуклый***, если каждая прямая, содержащая его сторону, является его опорной прямой. Это определение применяется также к многоугольнику с внутренностью.

Обобщая, вводят следующие

Определения. Кривая называется *выпуклой*, если в каждой ее точке у нее есть опорная прямая (рис. 88). Площадка, как и замкнутая область, называется *выпуклой*, если в каждой точке ее границы у нее есть опорная прямая (рис. 89).

Выпуклая ломаная и выпуклый многоугольник подпадают под эти определения, так как каждая точка ломаной принадлежит какому-нибудь ее звену.

Опорная плоскость. Выпуклая поверхность. Многогранник называется *выпуклым*, если он лежит по одну сторону от плоскости каждой своей грани (рис. 90). Это определение одинаково относится как к многограннику — фигуре, составленной из многоугольников, так и к телесному многограннику.

Как и для многоугольников, эти определения из школьного курса обобщаются.

Определение. Плоскость называется *опорной плоскостью* данной фигуры, если она имеет с фигурой или ее границей хотя бы одну общую точку, но вся фигура содержится в одном полупространстве,

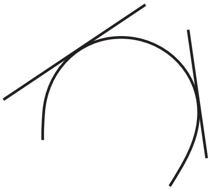


Рис. 88

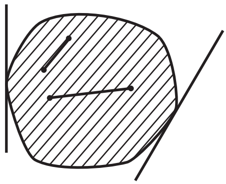


Рис. 89

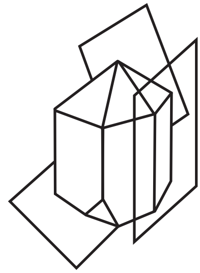


Рис. 90

ограниченном этой плоскостью (рис. 91). Говорят: плоскость *опорная* к фигуре F *в точке* A, если она при этом проходит через точку A фигуры F.

Пользуясь понятием опорной плоскости, можно пересказать определение выпуклого многогранника. Многогранник ***выпуклый***, если каждая плоскость, содержащая ее грань, является для него опорной. Обобщая, вводим

Определение. Поверхность называется *выпуклой*, если в каждой ее точке у нее есть опорная плоскость (рис. 92).

Выпуклый многогранник подпадает под это определение, так как каждая его точка принадлежит какой-нибудь его грани.

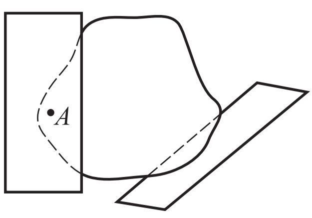


Рис. 91

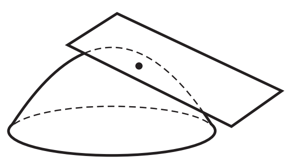


Рис. 92

Выпуклые фигуры. Есть другое понятие выпуклости, применяемое одинаково к фигурам как на плоскости, так и в пространстве.

Фигура называется *выпуклой*, если вместе с каждыми двумя своими точками она содержит и соединяющий их отрезок (на рис. 89 фигура выпуклая, на рис. 87 — не выпуклая).

Теорема 1. *Фигура на плоскости, обладающая внутренними точками и содержащая свою границу, выпукла тогда и только тогда,*

когда в каждой точке границы у нее есть опорная прямая (или, если фигура ограниченная, можно сказать, когда она ограничена выпуклой кривой) (рис. 89).

Совершенно такая же теорема, связывающая введенное здесь понятие выпуклой фигуры с предыдущим, выполняется в пространстве.

Теорема 2. *Фигура в пространстве, обладающая внутренними точками и содержащая свою границу, является выпуклой тогда и только тогда, когда в каждой точке границы у нее есть опорная плоскость* (или, если фигура ограниченная, можно сказать, когда она ограничена выпуклой поверхностью).

Выпуклость извне и внутри. Можно сказать, мы имеем два понятия выпуклости: внешнее и внутреннее. Фигура *выпукла извне*, если в каждой точке границы у нее на плоскости есть опорная прямая, а в пространстве — опорная плоскость. Это вполне согласуется с наглядным представлением о выпуклой фигуре и выпуклом теле.

С другой стороны, фигура *выпукла внутри*, если можно любые две ее точки соединить в ней отрезком. В случае, если у фигуры есть внутренние точки и она содержит свою границу, это равносильно тому, что из каждой внутренней точки можно дойти до любой точки границы, идя по прямой, как в комнате, где нет закоулков. Две сформулированные теоремы означают, что для содержащих свою границу фигур с внутренними точками оба понятия выпуклости равносильны. Такая фигура, выпуклая извне, выпукла также внутри, и обратно. Полное доказательство теорем 1, 2 не совсем просто, и мы не будем его излагать в полной общности, а ограничимся случаем многоугольников и многогранников.

Выпуклые многогранники и многоугольники. Рассматриваются плоские многоугольники и телесные многогранники.

Теорема 3. *Многоугольник или многогранник, выпуклый внутри, является выпуклым также извне.*

Доказательство. Пусть многогранник P — выпуклый внутри. Возьмем какую-либо его грань Q. Докажем, что он расположен по одну сторону от содержащей ее плоскости α.

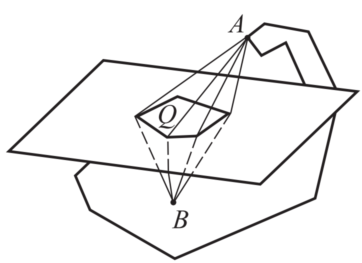


Рис. 93

Допустим противное: пусть у многогранника P есть точки A и B, лежащие по разные стороны от плоскости α (рис. 93). Тогда, поскольку многогранник P выпуклый внутри,

он должен содержать всякий отрезок, соединяющий A или B с точками грани Q, т. е. он должен содержать пирамиды с вершинами A, B и общим основанием Q. Но тогда многоугольник Q оказывается внутри многогранника и, значит, не может быть его гранью. Следовательно, многогранник лежит по одну сторону от плоскости грани, что и требовалось доказать.

Для многоугольника доказательство совершенно такое же. Разница лишь в том, что вместо грани берется сторона Q и строятся треугольники с вершинами A, B и основанием Q. □

Теорема 4. *Многоугольник или многогранник, выпуклый извне, является выпуклым также внутри.*

Мы докажем даже более общую теорему для любых площадок и тел, поскольку доказательство ничуть не сложнее.

Теорема 5. *Площадки и тела, выпуклые извне, являются выпуклыми и внутри.*

Доказательство. Пусть тело P выпукло извне, так что в каждой точке его границы есть опорная плоскость. Докажем, что тогда отрезок, соединяющий любую его внутреннюю точку A с какой-нибудь другой его точкой B, содержится в теле P. Действительно, допустим противное: пусть на отрезке AB есть точка C, не принадлежащая телу P. Тогда по теореме II.4.1 на этом отрезке есть точка границы тела. В этой точке, по условию, есть опорная плоскость. Эта плоскость α не может проходить через точку A, так как та лежит внутри P: иначе с обеих сторон от плоскости α лежали бы точки тела и плоскость α не была бы опорной (рис. 94). Следовательно, плоскость α пересекает отрезок AB. Но тогда точки A и B лежат с разных сторон от нее, а это противоречит тому, что они принадлежат телу и плоскость α — опорная к нему.

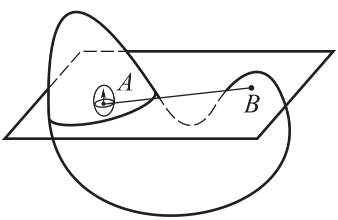


Рис. 94

Итак, отрезок, соединяющий две точки тела P, из которых хотя бы одна внутренняя, содержится в P.

Пусть теперь точки A, B лежат на границе тела. По самому определению тела его граница является границей его внутренности. Тем самым сколь угодно близко к граничным точкам A, B есть внутренние точки тела. Мы можем, стало быть, взять точки A_n, B_n $(n = 1,\ 2,\ \ldots)$ так, что $A_n \to A$, $B_n \to B$. Тогда отрезки A_nB_n «сходятся» к отрезку AB. Так как точки A_n, B_n внутренние, то, по доказанному, отрезки

A_nB_n содержатся в теле и даже внутри него. Поэтому и отрезок AB содержится в теле, что и требовалось доказать.

Доказательство теоремы для площадки то же самое. Нужно только в первой части доказательства, рассматривая отрезок AB, сослаться на существование опорной прямой; она разделит точки A и B. Читатель сам проведет это доказательство. □

О выпуклых фигурах. Речь идет о выпуклых фигурах в общем их определении — изнутри. Такая фигура на плоскости или в пространстве не обязана иметь внутренние точки. Пример представляют отрезок, прямая, одна точка, пустое множество. (По определению фигура выпукла, если для каждых двух ее точек отрезок, их соединяющий, содержится в фигуре; так что если двух точек нет, то условие выполнено.)

Теорема 6. *Пересечение любой совокупности выпуклых фигур есть выпуклая фигура.*

Доказательство. Точки A, B, принадлежащие пересечению фигур, принадлежат каждой из них. Поэтому если фигуры выпуклы, то и отрезок AB содержится в них всех и, значит, содержится в их пересечении. Значит, оно выпукло. □

Следствие 1. *Пересечение выпуклой фигуры с плоскостью есть выпуклая фигура.* □

Следствие 2. *Пересечение выпуклой фигуры с полупространством представляет собой выпуклую фигуру.* □

В этом смысле всякая плоскость, пересекающая выпуклую фигуру, делит ее на две выпуклые фигуры. (Точки исходной фигуры, лежащие на самой делящей плоскости, можно отнести к каждой из получающихся выпуклых фигур.)

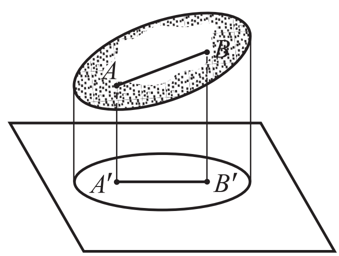


Рис. 95

Теорема 7. *Проекция пространственной выпуклой фигуры на плоскость есть выпуклая фигура* (это верно для каждой параллельной проекции).

Доказательство очевидно из рис. 95. □

Площадь выпуклой поверхности. Площадь многогранной поверхности естественно считать равной сумме площадей ее граней. Площадь поверхности выпуклого тела можно тогда определить как инфимум площадей заключающих это тело замкнутых многогранных

поверхностей. При этом достаточно ограничиться выпуклыми многогранными поверхностями, описанными вокруг исходного тела, т. е. такими, грани которых лежат в его опорных плоскостях.

Пользуясь таким определением, можно, например, показать, как это делается в школьном курсе, что площадь сферы радиуса R равна $4\pi R^2$. Для случая выпуклых поверхностей, образующих лишь часть границы выпуклого тела, наше определение требует некоторой модификации. Подробнее понятие площади поверхности (как и длины кривой) обсуждается в части 4. А в следующем параграфе мы узнаем, как найти площадь произвольного сферического многоугольника.

§ 4. Многогранные углы и сферические многоугольники

Связь многогранных углов и сферических многоугольников. Пусть V — многогранный угол с ребрами $a_1, \ldots, a_n$; они перенумерованы последовательно, так что угол V состоит из плоских углов $a_1a_2, a_2a_3, \ldots, a_na_1$. (Двугранный угол мы, как и раньше, не причисляем к многогранным углам.) Опишем вокруг его вершины единичную сферу. Всякий плоский угол с вершиной O пересекает ее по дуге большого круга с концами в точках, где сферу пересекают лучи — ребра угла. Таким образом, многогранному углу $a_1 \ldots a_n$ соответствует на сфере многоугольник $A_1 \ldots A_n$ (рис. 96).

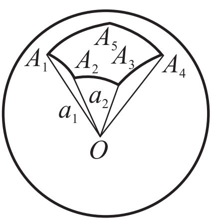


Рис. 96

И обратно: если на сфере дан многоугольник $A_1 \ldots A_n$, стороны которого — дуги больших кругов, то ему соответствует многогранный угол с вершиной в центре O сферы и с ребрами $OA_1, \ldots, OA_n$.

Если θ_k — угол (величина плоского угла a_ka_{k+1}) между ребрами a_k, a_{k+1} (полагаем $a_{n+1} = a_1$), то длина соответствующей стороны A_kA_{k+1} сферического многоугольника равна θ_kR, где R — радиус сферы. (Углы измеряем в радианах.)

Касательная плоскость к сфере перпендикулярна радиусу, проведенному в точку касания. Поэтому касательные к дугам — сторонам многоугольника $A_1 \ldots A_n$ — в любой его вершине перпендикулярны идущему в нее ребру многогранного угла. А угол между касательными — это, по определению, угол между дугами, к которым они проведены. Таким образом, углы сферического многоугольника $A_1 \ldots A_n$ равны двугранным углам многогранного угла $a_1 \ldots a_n$.

Так между многогранными углами и сферическими многоугольниками устанавливается полное соответствие: плоским углам соответствуют стороны, двугранным углам — углы между сторонами (при вершинах).

Определение. Многогранный угол называется *выпуклым*, если он расположен по одну сторону от плоскости каждой своей грани, т. е. в одном ограниченном ею полупространстве.

Нетрудно видеть, что среди плоских углов выпуклого многогранного угла нет развернутых и сверхтупых.

Лемма. *Через вершину выпуклого многогранного угла проходит такая плоскость, что весь он лежит с одной стороны от нее.*

Действительно, пусть V — какой-либо выпуклый многогранный угол. Тогда плоскости двух граней, смежных по произвольному его ребру a, ограничивают полупространства, содержащие многогранный угол V. Так что он содержится в пересечении этих полупространств. Это значит, что он содержится в двугранном угле W, ребром которого служит прямая $\overline{a}$, содержащая ребро a самого угла V. Луч a', дополнительный ребру a на этой прямой, не принадлежит углу V, так как тот не двугранный. Поэтому через вершину угла V можно провести такую плоскость, что луч a' проходит с одной стороны от нее, а все ребра угла V и, значит, сам угол V располагаются с другой стороны[16]. Эта плоскость не будет иметь с углом V ничего общего, кроме вершины; так что угол V за вычетом вершины содержится внутри ограниченного этой плоскостью полупространства. $\square$

Все это переносится на сферические многоугольники.

Определение. Сферический многоугольник называется *выпуклым*, если он лежит по одну сторону от каждой большой окружности, проходящей вдоль одной из его сторон (рис. 97).

Лемма. *Для выпуклого сферического многоугольника есть полусфера, внутри которой он содержится.*

Это полусфера, ограниченная окружностью, которую высекает на сфере плоскость, проходящая через вершину многогранного угла и не имеющая с ним других общих точек (рис. 98). $\square$

Замечание. У выпуклого многогранного угла каждый плоский угол меньше развернутого. Только такой угол может не иметь с

[16] Например, проведем биссекториальную полуплоскость α двугранного угла W и спроектируем на нее все ребра угла V. Это будут лучи, проведенные из вершины O угла V. Проведем из O в той же полуплоскости луч p, образующий с a' еще меньший угол. Тогда проходящая через луч p плоскость, перпендикулярная α, будет обладать указанным свойством.

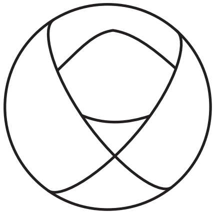


Рис. 97

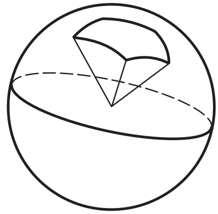


Рис. 98

плоскостью ничего общего, кроме вершины. Соответственно, стороны выпуклого многоугольника на сфере меньше большой полуокружности.

Определение. Для выпуклого многогранного угла естественно определяется соответствующий *телесный угол* как пересечение всех содержащих его полупространств, ограниченных плоскостями его плоских углов. Аналогично *выпуклый многоугольник с внутренностью* на сфере получается как пересечение полусфер, содержащих данный выпуклый многоугольник и ограниченных большими окружностями, содержащими его стороны (рис. 97).

Трехгранные углы и треугольники. Трехгранным углам соответствуют сферические треугольники. И те, и другие могут быть и не выпуклыми (рис. 99). Но мы будем рассматривать только выпуклые трехгранные углы и соответственно — только выпуклые сферические треугольники. У выпуклых трехгранных углов, как у всяких выпуклых многогранных углов, плоские углы меньше развернутого. Верно и обратное.

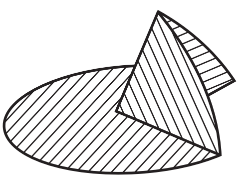


Рис. 99

Лемма. *Если у трехгранного угла все плоские углы меньше развернутого, то он выпуклый.*

Действительно, если через два ребра трехгранного угла провести плоскость α, то третье ребро оказывается с какой-то одной стороны от нее, и если плоские углы, которые оно ограничивает вместе с двумя другими ребрами, оба меньше развернутого угла, то и эти плоские углы лежат с той же стороны от плоскости α (рис. 100). Таким образом, трехгранный угол, у которого плоские углы меньше развернутого, располагается с одной стороны от плоскости каждой своей грани, т. е. он выпуклый. □

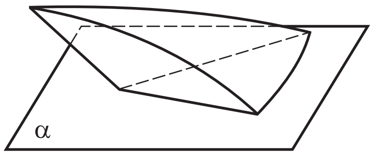


Рис. 100

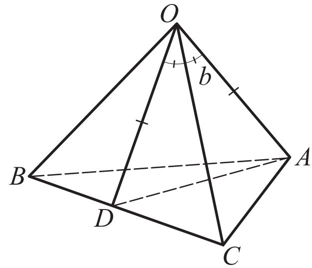


Рис. 101

Теорема 1. *У выпуклого трехгранного угла каждый плоский угол меньше суммы двух других. Соответственно: у выпуклого сферического треугольника каждая сторона меньше суммы двух других.*

Доказательство. Пусть имеется выпуклый трехгранный угол V с вершиной O, и пусть a, b, c — его плоские углы. Если все эти углы попарно равны, то утверждение теоремы очевидно выполнено.

Допустим, например, что угол a больше b. Возьмем на ребрах, противолежащих углам b, c, точки B, C и проведем отрезок BC. Получаем треугольник OBC. В нем $\angle O = a > b$. Поэтому можно провести в этом угле отрезок OD до стороны BC, образующий с OC угол, равный b (рис. 101).

На ребре, противолежащем углу a, возьмем такую точку A, что $OA = OD$. Тогда треугольники ODC и OAC равны. Поэтому $CD = CA$ и треугольник CAD равнобедренный, поэтому в нем угол D острый.

Следовательно, угол D в треугольнике ABD тупой, и потому противолежащая ему сторона бо́льшая. Так что $AB > BD$.

Теперь посмотрим на треугольники OAB и OBD. У них $OA = OD$ и сторона OB общая. Поэтому в том из них угол при O больше, у которого противолежащая сторона больше. И так как $AB > BD$, то, значит,

$$\angle AOB > \angle DOB.$$

Прибавляя к обеим частям этого неравенства равные углы $\angle DOC = \angle AOC = b$, получим

$$\angle AOB + \angle AOC > \angle BOC,$$

т. е. $c + b > a$, что и требовалось доказать. $\square$

Из доказанной теоремы 1 легко следует

Теорема 2. *Сумма плоских углов трехгранного (выпуклого) угла меньше* 2π.

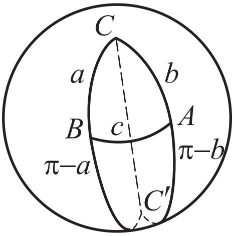


Рис. 102

Доказательство. Пусть a, b, c — плоские углы выпуклого трехгранного угла. Продолжив за вершину ребро, общее для углов a, b, получим трехгранный угол с плоскими углами $\pi - a$, $\pi - b$, смежными с a, b, и с третьим плоским углом c. На сфере это соответствует тому, что мы продолжаем стороны a, b сферического треугольника ABC и получаем сферический треугольник ABC' со сторонами $\pi - a$, $\pi - b$, c (рис. 102).

Применяя к полученному трехгранному углу (или, что равносильно, к треугольнику ABC') предыдущую теорему 1, можно написать

$$(\pi - a) + (\pi - b) > c.$$

Отсюда $a + b + c < 2\pi$, что и требовалось доказать. □

Доказанная теорема обобщается на любой выпуклый многогранный угол.

Теорема 3. *Сумма плоских углов выпуклого многогранного угла меньше* 2π.

Доказательство. Докажем теорему индукцией по числу граней угла. По теореме 2 данное утверждение верно для трехгранных углов. Допустим, что оно верно для n-гранных углов, и докажем его для $(n+1)$-гранных углов.

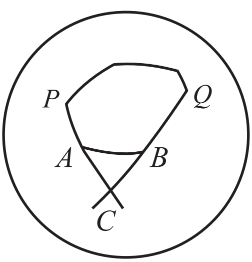


Рис. 103

Вместо многогранных углов можно рассматривать многоугольники на единичной сфере. Пусть имеется выпуклый сферический $(n+1)$-угольник. Возьмем какую-либо его сторону AB и продолжим соседние ей стороны AP, BQ за точки A, B до их пересечения; пусть C — точка пересечения (рис. 103). Получаем n-угольник со сторонами CP, CQ. Он выпуклый — как и тот многоугольник, из которого он получен (так же лежит по одну сторону от каждой большой окружности, содержащей его сторону).

По предположению сумма длин сторон этого многоугольника меньше 2π. Но она больше, чем у исходного, так как он получен из него заменой стороны AB на $AC + BC$.

Следовательно, и у исходного $(n+1)$-угольника сумма длин сторон меньше 2π. Теорема 3 доказана. □

Другое доказательство теоремы 3. Пусть дан выпуклый многогранный угол V с ребрами $a_1, \ldots, a_n$. Через его вершину проходит плоскость, не имеющая с ним других общих точек. Если ее сдвинуть параллельно в ту сторону, где расположен угол V, то она пересечет все его ребра и грани, и мы получим пирамиду с вершиной O и боковыми ребрами $OA_1, \ldots, OA_n$. У основания пирамиды n вершин $A_1, \ldots, A_n$, а потому сумма его углов α_i равна

$$\sum_{i=1}^{n} \alpha_i = (n-2)\pi. \tag{1}$$

При каждой вершине A_i грани образуют трехгранный угол; его плоский угол, принадлежащий основанию, меньше суммы двух других. Поэтому сумма углов основания меньше суммы всех углов β_i, γ_i боковых граней при каждой вершине A_i (рис. 104). То есть

$$\sum_{i=1}^{n} \alpha_i < \sum_{i=1}^{n} (\beta_i + \gamma_i). \tag{2}$$

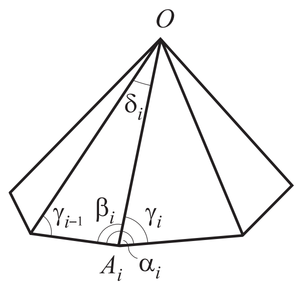


Рис. 104

Но $\sum\limits_{i=1}^{n} (\beta_i + \gamma_i)$ — это сумма всех углов всех n боковых граней без суммы их углов при вершине O, т. е.

$$\sum_{i=1}^{n} (\beta_i + \gamma_1) = n\pi - \sum_{i=1}^{n} \delta_i, \tag{3}$$

где $\sum\limits_{i=1}^{n} \delta_i$ — сумма всех углов при O.

Сопоставляя (1), (2), (3), получаем

$$\sum_{i=1}^{n} \alpha_i = (n-2)\pi < n\pi - \sum_{i=1}^{n} \delta_i,$$

откуда
$$\sum_{i=1}^{n} \delta_i < 2\pi,$$
что и требовалось доказать. □

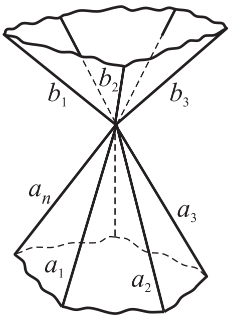


Рис. 105

Полярный многогранный угол. Пусть V — выпуклый многогранный угол, точка O — его вершина, лучи $a_1, \ldots, a_n$ — его ребра; $a_1a_2, a_2a_3, \ldots, a_na_1$ — его плоские углы. Проведем из O лучи $b_1, \ldots, b_n$, каждый из которых перпендикулярен соответствующей грани угла V:
$$b_1 \perp a_na_1,\ b_2 \perp a_1a_2,\ \ldots,\ b_n \perp a_{n-1}a_n. \quad (4)$$
Они ограничивают плоские углы b_1b_2, b_2b_3, $\ldots$, b_nb_1 (рис. 105). Эти плоские углы образуют многогранный угол W с ребрами $b_1, \ldots, b_n$. Этот угол называется *полярным* углу V. Так как $b_1 \perp a_na_1$, $b_2 \perp a_1a_2$, то ребро a_1 перпендикулярно b_1 и b_2 и тем самым перпендикулярно плоскости плоского угла b_1b_2: $a_1 \perp b_1b_2$. То же верно для ребра a_2 и угла b_2b_2 и т. д. Так что
$$a_1 \perp b_1b_2,\ a_2 \perp b_2b_3,\ \ldots,\ a_n \perp b_nb_1. \quad (5)$$
Это аналогично (4), и, значит, *угол V является полярным для угла W: отношение полярности симметрично.* Угол W также выпуклый (докажите).

Граням угла V отвечают перпендикулярные им ребра полярного угла W, и обратно.

Угол между перпендикулярами b_1, b_2 к граням a_na_1, a_1a_2 равен дополнению угла между этими гранями до развернутого, т. е. плоский угол b_1b_2 равен дополнению двугранного угла при ребре a_1 до π. И так в обе стороны (рис. 106; взгляд вдоль ребра a_1).

Итак, *если a — плоский угол одного из взаимно полярных углов и α — двугранный угол другого при перпендикулярном ребре, то*
$$a + \alpha = \pi.$$

Сумма углов сферического многоугольника. В случае плоскости сумма углов любого n-угольника равна $(n-2)\pi$. На сфере это уже не так.

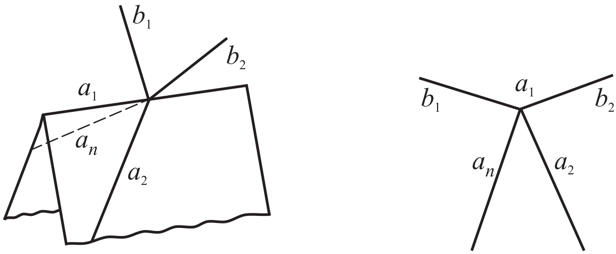


Рис. 106

Теорема 4. *Сумма двугранных углов (выпуклого) n-гранного угла больше $(n-2)\pi$.*

То же верно для углов сферического n-угольника, или, что равносильно: сумма углов, смежных с углами выпуклого сферического многоугольника, меньше 2π.

Доказательство. Ограничимся случаем выпуклого угла. Пусть V — выпуклый многогранный угол и $\alpha_1, \ldots, \alpha_n$ — его двугранные углы. Возьмем угол W, полярный V. Его плоские углы $a_1', \ldots, a_n'$ связаны с $\alpha_1, \ldots, \alpha_n$ так, что

$$a_1' = \pi - \alpha_1 \quad \text{и т. д.}$$

По теореме 3 $\sum\limits_{i=1}^{n} a_i' < 2\pi$. Поэтому

$$\sum_{i=1}^{n}(\pi - \alpha_i) < 2\pi \quad \text{и} \quad \sum_{i=1}^{n} \alpha_i > (n-2)\pi,$$

что и требовалось доказать. □

В частности, у трехгранного угла и соответственно у сферического треугольника

$$\alpha_1 + \alpha_2 + \alpha_3 > \pi.$$

Площадь сферического многоугольника. Следующая далее теорема 6 устанавливает, что суммы углов в теореме 4 больше $(n-2)\pi$ на площадь n-угольника.

Теорема 5. *Площадь выпуклого сферического треугольника, расположенного на единичной сфере, выражается через его углы следующим образом:*

$$S = \alpha + \beta + \gamma - \pi.$$

(На сфере радиуса R имеем $S=(\alpha+\beta+\gamma-\pi)R^2$.)

Доказательство. Пусть T — сферический треугольник с углами α, β, γ при вершинах A, B, C. Проведем большие окружности вдоль сторон AB, AC. Они пересекутся в точке A', симметричной A (относительно центра сферы), и ограничат два двуугольника с уг-

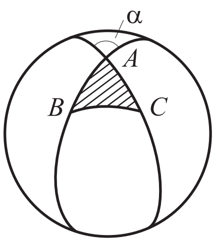


Рис. 107

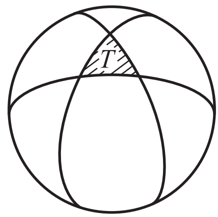


Рис. 108

лом α (рис. 107). Их площадь $S(D_A)$ составляет такую долю площади всей сферы, какую составляет 2α от полного угла вокруг A, т. е. от 2π:

$$S(D_A)=4\pi\frac{2\alpha}{2\pi}=4\alpha. \tag{6}$$

Совершенно так же, проводя большие окружности вдоль сторон BA, BC и CA, CB (рис. 108), получим фигуры D_B, D_C с площадями

$$S(D_B)=4\beta,\quad S(D_C)=4\gamma. \tag{7}$$

Пары проведенных больших окружностей, проходящих через A, B, C, пересекаются соответственно в точках A', B', C', симметричных точкам A, B, C. Так что получаются два симметричных треугольника $T=ABC$ и $T'=A'B'C'$, $S(T)=S(T')$.

Фигуры D_A, D_B, D_C покрывают всю сферу, но при этом покрывают треугольники T, T' трижды, т. е. два лишних раза. Поэтому

$$S(D_A)+S(D_B)+S(D_C)=4\pi+2S(T)+2S(T')=4\pi+4S(T).$$

Пользуясь (6) и (7), получаем

$$4(\alpha+\beta+\gamma)=4\pi+4S(T).$$

Отсюда

$$S(T) = \alpha + \beta + \gamma - \pi,$$

что и требовалось доказать. □

Теорема 6. *Площадь сферического n-угольника на единичной сфере выражается через его углы $\alpha_1, \ldots, \alpha_n$ формулой*

$$S = \sum \alpha_i - (n-2)\pi, \tag{8}$$

т. е. она равна разности суммы его углов и суммы углов плоского n-угольника.

Доказательство. Пусть P — выпуклый сферический многоугольник. Проводя диагонали из какой-нибудь его вершины, разобьем его на $(n-2)$ треугольника. Сумма их площадей даст площадь $S(P)$, а эта сумма равна сумме их углов минус $(\pi - 2)\pi$. Сумма же их углов равна сумме углов многоугольника P. Поэтому и получается (8).

Если многоугольник не выпуклый, то его, вообще говоря, нельзя разбить диагоналями из одной вершины. Но его вообще можно разбить на треугольники и получить ту же формулу (8). Но проводить этот вывод мы не будем. □

§ 5. Тригонометрия трехгранных углов и сферических треугольников

Мы рассматриваем выпуклые трехгранные углы и, соответственно, — выпуклые сферические треугольники, опуская указание на их выпуклость, — она будет подразумеваться.

Если a — длина стороны треугольника на сфере радиуса R, то соответствующий плоский угол (в радианной мере) равен a/R. Но мы будем рассматривать треугольники на единичной сфере — с $R = 1$. Тогда a одинаково обозначает и длину стороны, и меру угла. Так будем понимать $\sin a$ и т. п. Для того чтобы перейти к треугольникам на сфере любого радиуса R, нужно брать $\sin \frac{a}{R}$ и т. п.

Трехгранный угол с прямым двугранным углом. Прямоугольный сферический треугольник. Сформулируем первое из двух равносильных друг другу утверждений.

Теорема 1. *Пусть a, b — катеты и c — гипотенуза прямоугольного сферического треугольника, α — угол, противолежащий катету a. Выполняются три соотношения:*

$$\cos c = \cos a \, \cos b \text{ — теорема косинусов,} \tag{1}$$

$$\sin a = \sin c \, \sin\alpha \text{ — теорема синусов,} \tag{2}$$

$$\operatorname{tg} b = \operatorname{tg} c \, \cos\alpha \text{ — теорема тангенсов.} \tag{3}$$

Аналогично $\sin b = \sin c \, \sin\beta$, $\operatorname{tg} a = \operatorname{tg} c \, \cos\beta$ *и т. д.*

Этой теореме соответствует теорема о трехгранных углах.

Теорема 1а. *Пусть в трехгранном угле один двугранный угол прямой, и пусть a, b — образующие его плоские углы, c — ему противолежащий и α — двугранный угол, противолежащий a. Тогда для этих углов выполняются соотношения* (1), (2), (3).

Доказательство. Пусть дан трехгранный угол с такими углами, как сказано в теореме. Для определенности будем считать угол α острым (это ничего не меняет). Пусть O — вершина и B — какая-нибудь точка на ребре угла β, противолежащего b.

Опустим из точки B перпендикуляр BC на плоскость угла b. Так как плоскость угла a перпендикулярна плоскости угла b, то перпендикуляр BC лежит в плоскости угла a и перпендикулярен плоскости угла b (рис. 109).

Если теперь CA — перпендикуляр к ребру угла α, то отрезок BA также перпендикулярен этому ребру — по теореме II.3.8 «о трех перпендикулярах». Так как BC перпендикулярен плоскости угла b, то $BC \perp AC$.

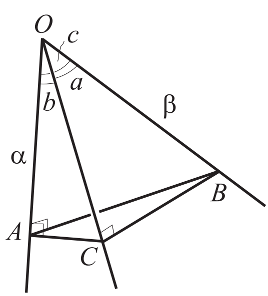


Рис. 109

Итак, мы имеем прямоугольный треугольник ABC и еще три прямоугольных треугольника OBC, OAC, OBA. У первого угол C прямой, у двух других угол A прямой.

У треугольника OBC угол O — это угол a, поэтому

$$OC = OB \, \cos a.$$

У треугольника OAC угол при O — это угол b, и угол A — прямой. Поэтому

$$OA = OC \cdot \cos b,$$

и, применяя предыдущее равенство, получаем

$$OA = OB \, \cos a \, \cos b.$$

У треугольника OBA угол при O — это угол c, и угол A — прямой. Поэтому

$$OA = OB \, \cos c.$$

Сравнивая с предыдущим равенством, получаем

$$\cos c = \cos a \, \cos b.$$

Таким образом, равенство (1) — теорема косинусов — доказано.

Для доказательства двух других соотношений (2), (3) обратимся к треугольнику ABC. В нем угол C прямой, стороны — отрезки AB, AC — перпендикулярны ребру OA. Поэтому угол A — это угол α при этом ребре и

$$BC = AB \, \sin \alpha, \quad AC = AB \, \cos \alpha. \tag{4}$$

Далее, отрезок BC — это катет в прямоугольном треугольнике OBC, противолежащий углу a.

Аналогично AB — катет в треугольнике OAB, противолежащий углу c. Поэтому

$$BC = OB \, \sin a, \quad AB = OB \, \sin c.$$

Подставляя эти выражения в первое из равенств (4), получим

$$\sin a = \sin c \, \sin \alpha.$$

Это теорема синусов (2).

Теперь обратимся к треугольникам OAB и OAC. В первом из них угол при A прямой, а угол при O равен c. Поэтому

$$AB = OA \operatorname{tg} c.$$

В треугольнике OAC угол A прямой, а угол при O равен b. Поэтому

$$AC = OA \operatorname{tg} b.$$

Подставляя эти два равенства во второе равенство (4) и сокращая на OA, получаем

$$\operatorname{tg} b = \operatorname{tg} c \, \cos \alpha.$$

Таким образом, соотношение (3) тоже доказано. □

Соотношения в произвольном выпуклом сферическом треугольнике или трехгранном угле.

Теорема 2. *Пусть a, b, c — стороны, α, β, γ — противолежащие им углы сферического треугольника. Выполняются соотношения:*

$$\frac{\sin \alpha}{\sin a} = \frac{\sin \beta}{\sin b} = \frac{\sin \gamma}{\sin c} \text{ — теорема синусов,}$$

$$\cos c = \cos a \cos b + \sin a \sin b \cos\gamma \text{ — теорема косинусов}$$

и аналогично для $\cos a$, $\cos b$.

Этому соответствует теорема о трехгранных углах.

Теорема 2а. *Для плоских углов* a, b, c *и противолежащих им двугранных углов* α, β, γ *трехгранного угла выполняются указанные в теореме* 2 *соотношения — теоремы синусов и косинусов.*

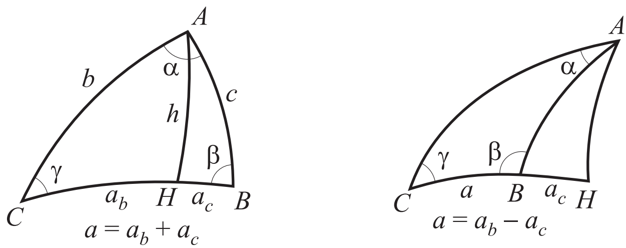


Рис. 110

Д о к а з а т е л ь с т в о. Обе теоремы — синусов и косинусов — доказываются для сферических треугольников совершенно аналогично тому, как это делается для плоских треугольников при помощи результатов для прямоугольных треугольников.

Пусть дан сферический треугольник ABC, в нем $\angle A = \alpha$, $\angle B = \beta$, $\angle C = \gamma$ $(\beta,\ \gamma \neq 90^\circ)$ и $BC = a$, $CA = b$, $AB = c$, см. рис. 110.

Проведем из вершины A высоту AN.[17] Получаем два прямоугольных треугольника AHB, AHC. Полагая $AH = h$ и пользуясь теоремой синусов, получим

$$\sin h = \sin c \sin\beta, \quad \sin h = \sin b \sin\gamma.$$

Отсюда

$$\frac{\sin\beta}{\sin b} = \frac{\sin\gamma}{\sin c}.$$

Аналогично, проводя высоту из вершины B, получим, что $\sin\alpha/\sin a = \sin\gamma/\sin c$. Поэтому

[17] В трехгранном угле с вершиной O этому соответствует плоскость, проходящая через ребро OA перпендикулярно грани OBC.

$$\frac{\sin\alpha}{\sin a} = \frac{\sin\beta}{\sin b} = \frac{\sin\gamma}{\sin c},$$
и теорема синусов доказана.

Докажем теорему косинусов. С этой целью применим теорему косинусов для прямоугольных треугольников к $\triangle AHC$. Тогда, полагая $BH = a_c$, получим

$$\cos c = \cos a_c \, \cos h. \tag{5}$$

Если точка H лежит на BC, то $BC = BH + HC$. Если же B на HC, то $BC = HC - BH$ (рис. 110). Полагая $CH = a_b$, имеем:

$$\text{либо} \quad a = a_b + a_c, \quad \text{либо} \quad a = a_b - a_c,$$

так что

$$a_c = \pm(a - a_b). \tag{6}$$

Кроме того, из треугольника AHC видно, что $\cos b = \cos a_b \, \cos h$, т. е.

$$\cos h = \frac{\cos b}{\cos a_b}. \tag{7}$$

Теперь, пользуясь (6) и (7), получим из (5) (поскольку косинус сохраняется при перемене знака угла)

$$\cos c = \cos{(a - a_b)}\frac{\cos b}{\cos a_b}, \tag{8}$$

а так как $\cos{(a - a_b)} = \cos a \, \cos a_b + \sin a \, \sin a_b$, то

$$\cos c = \cos a \, \cos b + \sin a \, \sin a_b \frac{\cos b}{\cos a_b}. \tag{9}$$

Применив теорему тангенсов для прямоугольных треугольников к $\triangle AHC$, получим:

$$\operatorname{tg} a_b = \operatorname{tg} b \, \cos\gamma.$$

Подставляя это в (9), получаем

$$\cos c = \cos a \, \cos b + \sin a \, \sin b \, \cos\gamma.$$

Мы предполагали, что либо точка H лежит на стороне BC, либо B — на HC. Но есть еще возможность, что C лежит на BH. Тогда $BH = BC + CH$, т. е.

$$a_c = a + a_b.$$

В таком случае в формуле (8) вместо $\cos(a-a_b)$ будет стоять $\cos(a+a_b)$, и поэтому в (9) второй член будет с минусом. Но в этом случае в треугольнике AHC угол при C равен $\pi-\gamma$, так что $\operatorname{tg} a_b = \operatorname{tg} b \cos(\pi-\gamma) = -\operatorname{tg} b \cos\gamma$.

Таким образом, «минус на минус дает плюс», и получается та же формула (9) теоремы косинусов. $\square$

Другое доказательство теоремы косинусов.

Пусть дан трехгранный угол с вершиной O и ребрами OA, OB, OC. Пусть a, b, c — его плоские углы, противолежащие соответственно ребрам OA, OB, OC, а γ — двугранный угол при ребре OC (рис. 111). Считаем a, $b \neq 90°$.

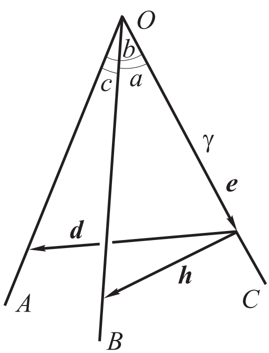


Рис. 111

Возьмем на ребре OC единичный вектор $\boldsymbol{e}$, и пусть $\boldsymbol{d}$, $\boldsymbol{h}$ — такие перпендикулярные ему векторы, что $\boldsymbol{e}+\boldsymbol{d}$ оказывается на ребре OA, а $\boldsymbol{e}+\boldsymbol{h}$ — на OB.

Вычислим скалярные произведения этих векторов на ребрах. Так как $\boldsymbol{e}^2 = 1$ и $\boldsymbol{e} \perp \boldsymbol{d}$, то $\boldsymbol{e}(\boldsymbol{e}+\boldsymbol{d}) = 1$.

С другой стороны, угол между $\boldsymbol{e}$ и $\boldsymbol{e}+\boldsymbol{d}$ — это угол между ребрами OC и OA, т. е. угол b. Поэтому

$$\boldsymbol{e}(\boldsymbol{e}+\boldsymbol{d}) = |\boldsymbol{e}||\boldsymbol{e}+\boldsymbol{d}|\cos b = |\boldsymbol{e}+\boldsymbol{d}|\cos b.$$

Вместе с предыдущим равенством это дает

$$|\boldsymbol{e}+\boldsymbol{d}|\cos b = 1, \qquad \frac{1}{|\boldsymbol{e}+\boldsymbol{d}|} = \cos b. \tag{10}$$

Совершенно аналогично, беря $\boldsymbol{e}(\boldsymbol{e}+\boldsymbol{h})$, выводим, что

$$|\boldsymbol{e}+\boldsymbol{h}|\cos a = 1, \qquad \frac{1}{|\boldsymbol{e}+\boldsymbol{h}|} = \cos a. \tag{11}$$

Теперь вычисляем:

$$(\boldsymbol{e}+\boldsymbol{d})(\boldsymbol{e}+\boldsymbol{h}) = 1 + \boldsymbol{dh}$$

(поскольку $\boldsymbol{e}^2 = 1$ и $\boldsymbol{e} \perp \boldsymbol{d}$ и $\boldsymbol{h}$, т. е. $\boldsymbol{ed} = \boldsymbol{eh} = 0$).

С другой стороны, угол между $(\boldsymbol{e}+\boldsymbol{d})$ и $(\boldsymbol{e}+\boldsymbol{h})$ — это угол между ребрами OA, OB, т. е. угол c.

Вместе с предыдущим выражением это дает

$$|\boldsymbol{e}+\boldsymbol{d}||\boldsymbol{e}+\boldsymbol{h}|\cos c = 1+\boldsymbol{dh}.$$

Отсюда, пользуясь формулами (11), (10), получаем

$$\cos c = \cos a\ \cos b + \frac{\boldsymbol{dh}}{|\boldsymbol{e}+\boldsymbol{d}||\boldsymbol{e}+\boldsymbol{h}|}. \tag{12}$$

Вычислим второй член этой формулы. Вспомним, что векторы $\boldsymbol{d}$, $\boldsymbol{h}$ перпендикулярны ребру OC. Поэтому угол между ними равен двугранному углу γ при этом ребре. Следовательно,

$$\boldsymbol{dh} = |\boldsymbol{d}|\cdot|\boldsymbol{h}|\cdot\cos\gamma. \tag{13}$$

Далее, так как $\boldsymbol{d} \perp \boldsymbol{e}$, то $(\boldsymbol{e}+\boldsymbol{d})$ — гипотенуза в прямоугольном треугольнике с катетами $\boldsymbol{d}$, $\boldsymbol{e}$. По построению вектор $\boldsymbol{e}$ лежит на ребре OC, $\boldsymbol{e}+\boldsymbol{d}$ — на ребре OA. Поэтому угол, противолежащий катету $\boldsymbol{d}$, — это угол b. Таким образом,

$$|\boldsymbol{d}| = |\boldsymbol{e}+\boldsymbol{d}|\sin b. \tag{14}$$

Совершенно аналогично выводим, что

$$|\boldsymbol{h}| = |\boldsymbol{e}+\boldsymbol{h}|\sin a. \tag{15}$$

Теперь, пользуясь (12)–(14), получаем

$$\frac{\boldsymbol{dh}}{|\boldsymbol{e}+\boldsymbol{d}||\boldsymbol{e}+\boldsymbol{h}|} = \sin a\ \sin b\ \cos\gamma.$$

Подставляя это в (12), получаем

$$\cos c = \cos a\ \cos b + \sin a\ \sin b\ \cos\gamma. \quad \square$$

Теорема 3 (вторая теорема косинусов). *Для трехгранного угла (или сферического треугольника) при тех же обозначениях, что в теоремах* 2, 2а, *выполняется равенство*

$$\cos\gamma = -\cos\alpha\ \cos\beta + \sin\alpha\ \sin\beta\ \cos c \tag{16}$$

и аналогично для $\cos\alpha$, $\cos\beta$.

Доказательство. Для данного трехгранного угла с плоскими углами a, b, c и двугранными углами α, β, γ построим полярный ему

трехгранный угол. У него плоские углы будут $a' = \pi - \alpha$, $b' = \pi - \beta$, $c' = \pi - \gamma$ и двугранный угол $\gamma' = \pi - c$. По теореме косинусов

$$\cos c' = \cos a' \, \cos b' + \sin a' \, \sin b' \, \cos \gamma'.$$

Это и даст (16). □

Выводы из первой и второй теорем косинусов. В качестве следствий теорем косинусов мы получим здесь теоремы, уже доказанные в предыдущем параграфе независимым путем (см. теоремы 1, 2, 3, § 4). Однако проследить имеющиеся здесь связи будет поучительно. Теоремы 4, 5 мы формулируем для сферических треугольников, опуская уже неоднократно повторявшийся пересказ для трехгранных углов.

Теорема 4. *В треугольнике сумма двух сторон больше третьей.*

Доказательство. По теореме косинусов

$$\cos c = \cos a \, \cos b + \sin a \, \sin b \, \cos \gamma.$$

Поскольку заведомо $\cos \gamma > -1$, то

$$\cos c > \cos a \, \cos b - \sin a \, \sin b = \cos (a + b).$$

И так как косинус — убывающая функция угла, то $c < a + b$, что и требовалось доказать. □

Теорема 5. *Сумма углов треугольника больше* π:

$$\alpha + \beta + \gamma > \pi.$$

Доказательство. Так как $\cos c < 1$, то из второй теоремы косинусов

$$\cos \gamma < -\cos \alpha \, \cos \beta + \sin \alpha \, \sin \beta = -\cos (\alpha + \beta).$$

Или так как $\cos \gamma = -\cos (\pi - \gamma)$, то $\cos (\pi - \gamma) > \cos (\alpha + \beta)$. И так как косинус — убывающая функция угла, то $\pi - \gamma < \alpha + \beta$ и $\alpha + \beta + \gamma > \pi$. □

Теорема 6. *Сумма плоских углов трехгранного угла меньше* 2π.

Доказательство. У трехгранного угла, полярного данному, двугранные углы будут $\alpha' = \pi - a$, $\beta' = \pi - b$, $\gamma' = \pi - c$, и так как по предыдущему $\alpha' + \beta' + \gamma' > \pi$, то $3\pi - (a + b + c) > \pi$, т. е.

$$a + b + c < 2\pi. \quad \square$$

Замечание. (Переход к плоскости.) При малом x

$$\sin x = x(1+\varepsilon), \quad \cos x = 1 - \frac{x^2}{2}(1+\delta),$$

где ε и δ — бесконечно малые вместе с x (обе порядка x^2). Поэтому для малых сферических треугольников получаем из теоремы синусов

$$\frac{\sin\alpha}{\sin\beta} = \frac{a}{b}(1+\xi)$$

из теоремы косинусов —

$$\frac{1}{c^2}(a^2 + b^2 - 2ab\,\cos\gamma) = 1 + \eta$$

с бесконечно малыми ξ, η. Поэтому в пределе получаются теоремы синусов и косинусов для плоских треугольников.

Если c мало, то $\cos c = 1 - \varepsilon$, где ε стремится к нулю вместе с $c(\varepsilon \sim c^2)$. Поэтому из второй теоремы косинусов

$$\cos\gamma = -\cos\alpha\,\cos\beta + \sin\alpha\,\sin\beta - \delta = -\cos(\alpha+\beta) - \delta.$$

Это равенство означает, что $\alpha+\beta+\gamma$ близко к π, и что при $c \to 0$, когда $\delta \to 0$, $\alpha+\beta+\gamma \to \pi$. Вторая теорема косинусов в пределе переходит в теорему о сумме углов треугольника.

Часть 3
ПРЕОБРАЗОВАНИЯ. ДРУГИЕ ГЕОМЕТРИИ

Основной предмет третьей части — различные классы (точнее, группы) преобразований фигур в евклидовом пространстве. С наибольшей полнотой в главе I изучены важнейшие преобразования — наложения, т. е. преобразования, сохраняющие расстояния (их также называют движениями или перемещениями). С наложениями тесно связано понятие о симметрии фигур, их «правильности».

В последующих трех главах классы преобразований последовательно расширяются: подобия сохраняют отношение отрезков, аффинные преобразования — отношения параллельных отрезков, а проективные преобразования — лишь прямолинейность расположения точек.

Каждой группе преобразований отвечает совокупность свойств фигур, сохраняющихся при этих преобразованиях. Выделение этих свойств — аффинных, проективных — приводит к «другим геометриям» — аффинной и проективной. Именно таким путем (а не аксиоматическим) мы излагаем основные положения этих геометрий.

В последней главе этой части, расширяя аксиоматику евклидовой геометрии, мы строим основы многомерной евклидовой геометрии.

Глава I
НАЛОЖЕНИЯ

§ 1. Отдельные виды наложений

Параллельный перенос. Он определяется одинаково на плоскости и в пространстве. *Параллельным переносом*, или, короче, *переносом* (или еще *сдвигом*), фигуры называется такое отображение, при котором точки смещаются в одном и том же направлении на равные расстояния, т. е. при переносе фигуры каждым двум точкам X, Y со-

поставляются такие точки X', Y', что (рис. 1)

$$\overrightarrow{XX'} = \overrightarrow{YY'}. \tag{1}$$

Это равенство означает, что направленные отрезки $\overrightarrow{XX'}$, $\overrightarrow{YY'}$ представляют один и тот же вектор. Стало быть, перенос — это отображение, при котором все точки смещаются на один и тот же вектор — «вектор переноса». Так что если этот вектор обозначить $\boldsymbol{t}$, то для всех точек фигуры $\overrightarrow{XX'} = \boldsymbol{t}$.

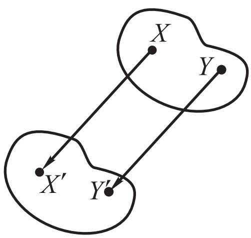


Рис. 1

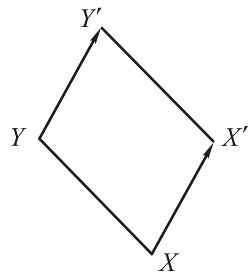


Рис. 2

Теорема 1. *Параллельный перенос характеризуется тем, что он сохраняет расстояние и направление, т. е. каждым двум точкам X, Y сопоставляет такие точки X', Y',* что (рис. 2)

$$\overrightarrow{X'Y'} = \overrightarrow{XY}. \tag{2}$$

(Так как отображение, сохраняющее расстояние, — это наложение, то перенос можно определить как наложение, сохраняющее направление.)

Доказательство. Сказанное следует из того, что равенства (1) и (2) равносильны. □

Композиция параллельных переносов есть параллельный перенос, и отображение, обратное параллельному переносу, есть параллельный перенос.

Это непосредственно следует из того, что перенос — это отображение, сохраняющее расстояние и направление.

Параллельный перенос задается одной парой соответствующих точек: если указано, какая точка A' соответствует данной точке A, то

какая точка X' отвечает любой другой точке X, определяется равенством

$$\overrightarrow{XX'} = \overrightarrow{AA'}.$$

Пара точек A, A' задает вектор переноса $\boldsymbol{a} = \overrightarrow{AA'}$. При переносе на вектор $\boldsymbol{a}$ радиус-вектор $\overrightarrow{OX}$ произвольной точки X получает слагаемое $\boldsymbol{a}$:

$$\overrightarrow{OX'} = \vec{O}X + \boldsymbol{a}.$$

Поэтому если x, y, z — координаты точки X, а x', y', z' — точки X'; a, b, c — координаты вектора $\boldsymbol{a}$, то

$$x' = x + a, \quad y' = y + b, \quad z' = z + c.$$

Так представляется перенос в координатах в пространстве. На плоскости — так же с двумя координатами.

Перенос любой фигуры можно распространить на все пространство, перенося любую точку X на вектор переноса $\boldsymbol{a}$.

Между векторами и переносами есть полное соответствие: 1) каждый вектор определяет перенос, и обратно: каждый перенос задается вектором; 2) сложению векторов соответствует композиция переносов (она представляется сложением векторов этих переносов); 3) противоположный вектор соответствует обратному переносу.

Сказанное требует уточнения. Перенос, как всякое отображение, определяется парами соответственных точек. Поэтому переносы разных фигур — это разные отображения, если они задаются одним и тем же вектором, но на разных фигурах. Перенос точки A в точку A' — не то же, что перенос точки B — в B', хотя векторно может быть $\overrightarrow{AA'} = \overrightarrow{BB'}$. Поэтому точное соответствие между векторами и переносами есть только тогда, когда берутся переносы всего пространства или плоскости, если имеются в виду векторы на плоскости.

Центральная симметрия. Центральная симметрия определяется одинаково и на плоскости, и в пространстве.

Точки A, A' называются ***симметричными*** *относительно точки* O, если она служит серединой отрезка AA' (рис. 3). Точка O считается симметричной сама себе (относительно O).

Две *фигуры* называются *симметричными относительно точки* O, если они образованы попарно симметричными точками, т. е. для каждой точки одной фигуры есть точка, симметричная ей относительно O в другой фигуре, и обратно (рис. 4).

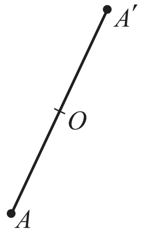


Рис. 3

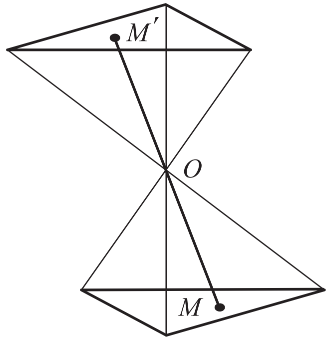


Рис. 4

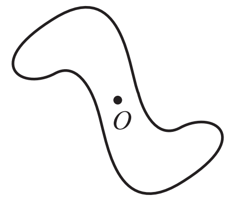


Рис. 5

В частности, фигура может быть симметрична сама себе. Тогда для каждой ее точки в ней есть точка, симметричная относительно O. В таком случае точка O называется *центром симметрии* фигуры, а сама фигура — *центрально симметричной* (*относительно* O) (рис. 5). Примеры: шар симметричен относительно своего центра, параллелепипед — относительно точки пересечения диагоналей, круговой цилиндр — относительно середины своей оси, и т. д.

Центральной симметрией с центром O называется отображение, сопоставляющее каждой точке отображаемой фигуры точку, ей симметричную относительно O.

Отношение между симметричными точками взаимно: если A' симметрична A, то и A симметрична A'; поэтому отображение, обратное центральной симметрии, — это та же центральная симметрия.

Теорема 2. *Центральная симметрия сохраняет расстояния, а все направления изменяет на противоположные.* То есть если при центральной симметрии точкам X, Y соответствуют точки X', Y', то

$$\overrightarrow{X'Y'} = -\overrightarrow{XY}.$$

Обратно: *отображение с такими свойствами есть центральная симметрия.*

Доказательство. Пусть при симметрии с центром O точки X, Y отображаются в X', Y'. Тогда, как следует из определения центральной симметрии,

$$\overrightarrow{OX'} = -\overrightarrow{OX}, \quad \overrightarrow{OY'} = -\overrightarrow{OY}. \tag{3}$$

Вместе с тем $\overrightarrow{XY} = \overrightarrow{OY} - \overrightarrow{OX}$, $\overrightarrow{X'Y'} = \overrightarrow{OY'} - \overrightarrow{OX'}$. Поэтому из (3)

следует

$$\overrightarrow{X'Y'} = -\overrightarrow{OY} + \overrightarrow{OX} = -\overrightarrow{XY}.$$

Обратно: пусть имеется отображение фигуры F, сохраняющее расстояние и изменяющее направление на противоположное. Возьмем какую-либо точку $A \in F$, и пусть $A' = f(A)$ — ее образ, O — середина отрезка AA' (или сама точка A, если A' совпадает с A). Возьмем любую точку $X \in F$ и ее образ $X' = f(X)$. По условию расстояние AX сохраняется, а направление $\overrightarrow{AX}$ изменяется на противоположное, поэтому

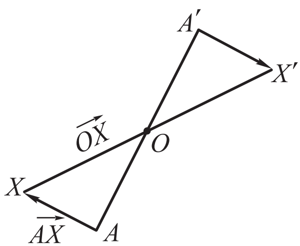


Рис. 6

$$\overrightarrow{A'X'} = -\overrightarrow{AX}. \tag{4}$$

Вместе с тем, поскольку O — середина AA', то (рис. 6)

$$\overrightarrow{OA'} = -\overrightarrow{OA}. \tag{5}$$

Кроме того,

$$\overrightarrow{OX} = \overrightarrow{OA} + \overrightarrow{AX}, \quad \overleftarrow{OX'} = \overrightarrow{OA'} + \overrightarrow{A'X'}.$$

Из этих равенств в силу(4) и (5) следует

$$\overrightarrow{OX'} = -\overrightarrow{OX}.$$

Стало быть, точка X' симметрична X относительно точки O; и так как это верно для любой точки X данной фигуры, то рассматриваемое отображение f есть ее центральная симметрия, что и требовалось дока-
зать. □

Центральная симметрия задается одной парой соответствующих точек: если A отображается на A', то центр — это середина отрезка AA'.

Так как центральная симметрия сохраняет расстояния, то она представляет собою наложение, и доказанную теорему можно выразить так:

Центральная симметрия — это наложение, изменяющее все направления на противоположные.

Центральная симметрия любой фигуры естественно распространяется на все пространство. Каждой точке сопоставляется точка, ей симметричная относительно того же центра. Так же распространяется на всю плоскость центральная симметрия плоской фигуры.

Но между центральной симметрией на плоскости и в пространстве есть очень существенная разница.

На плоскости центральная симметрия представляет собой поворот на 180° вокруг центра симметрии, как это непосредственно видно, но в пространстве она к повороту не сводится.

Отражение (зеркальная симметрия). Скользящее отражение. На плоскости зеркальную симметрию представляет отражение относительно прямой, в пространстве ее представляет отражение относительно плоскости.

Точки A, A' называются *симметричными относительно плоскости* α, если отрезок AA' перпендикулярен этой плоскости и делится ею пополам (рис. 7). Любая точка плоскости α считается симметричной сама себе относительно этой плоскости.

Совершенно так же определяются точки, симметричные относительно какой-либо прямой: нужно только в данном определении заменить плоскость на прямую (рис. 8). Дальше мы будем формулировать определения и результаты для отражения относительно плоскости; для отражения относительно прямой на плоскости они совершенно аналогичны.

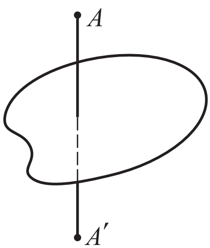


Рис. 7

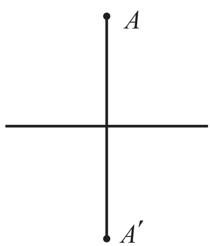


Рис. 8

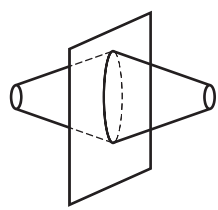


Рис. 9

Две *фигуры* называются *симметричными относительно плоскости*, если их точки попарно симметричны относительно этой плоскости, т. е. каждой точке одной фигуры отвечает симметричная ей точка другой фигуры, и обратно. Если эти фигуры совпадают, т. е. если это одна фигура, то говорят, что она *симметрична* относительно данной плоскости и что эта плоскость является ее *плоскостью симметрии* (рис. 9).

Отображение фигуры, при котором каждой ее точке соответствует точка, симметричная ей относительно данной плоскости, называется

отражением фигуры относительно этой плоскости (или *симметрией относительно этой плоскости*).

Симметричность точек очевидно взаимна: если A' симметрична A, то A симметрична A'. Поэтому отражение самому себе обратно.

Теорема 3. *Отражение сохраняет расстояния.*

Доказательство. Пусть отражение происходит в плоскости α. Введем прямоугольные координаты так, чтобы α была плоскостью $z = 0$. Тогда отражение представляется как перемена знака z: точке $A(x,\ y,\ z)$ сопоставляется точка $A'(x,\ y,\ -z)$. В выражение расстояния входит квадрат разности координат; он не изменяется при перемене их знака. Поэтому и расстояние не изменяется, что и требовалось доказать.

В случае отражения относительно прямой a на плоскости вводим координаты, при которых эта a была бы прямой $y = 0$ (т. е. осью x), и повторяем предыдущее рассуждение. □

Отражение задается парой соответственных точек. Плоскость отражения проходит через середину соединяющего их отрезка перпендикулярно ему.

Отражение фигуры в данной плоскости очевидным образом распространяется на все пространство.

Теорема 4. *Наложение пространства, при котором все точки некоторой плоскости неподвижны, является либо отражением относительно этой плоскости, либо тождественным отображением.*

Доказательство. Наложение отображает прямые на прямые и сохраняет углы. Поэтому если все точки плоскости неподвижны, то прямые, ей перпендикулярные, отображаются на себя. Отсюда следует сказанное. (Как?) □

Совершенно так же наложение плоскости, при котором все точки некоторой прямой неподвижны, является либо отражением относительно этой прямой, либо тождественным отображением.

Скользящим отражением называется композиция отражения в плоскости с переносом, параллельным этой плоскости. На плоскости скользящее отражение — это композиция отражения относительно прямой и переноса вдоль этой прямой. Порядок, в каком производятся отражение и перенос, безразличен: эти отображения перестановочны (как легко непосредственно убедиться).

Симметрия (отражение) относительно прямой в пространстве. Это отображение определяется дословно так же, как отражение относительно прямой на плоскости. Точка A' симметрична на точке A относительно прямой a, если отрезок AA' перпендикулярен прямой a

и делится ею пополам (рис. 10). Точки прямой a симметричны сами себе. Отражение в прямой a состоит в том, что каждой точке A ставится в соответствие точка A', симметричная A относительно прямой a. Это отображение сохраняет расстояния. Для доказательства введем координаты с осью z по прямой a. Тогда отражение в прямой a изобразится переменой знаков у x и y: точки (x, y, z) переходят в $(-x, -y, z)$, и из формулы для расстояния следует, что оно сохраняется.

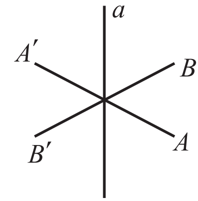


Рис. 10

В каждой плоскости α, перпендикулярной оси a, происходит центральная симметрия относительно точки пересечения a и α. Поэтому как центральная симметрия в плоскости есть поворот на $180°$ вокруг центра, так симметрия относительно прямой в пространстве есть поворот вокруг нее на $180°$.

§ 2. Повороты

Поворот в плоскости. Выражая наглядное представление о повороте вокруг точки, его определяют следующим образом. Поворот фигуры F вокруг точки O состоит в том, что ее точкам X сопоставляются точки X' так, что

1) отрезки OX, OX' равны;

2) углы между отрезками OX, OX' равны для всех точек X и откладываются от OX к OX' в одну и ту же сторону (по часовой стрелке, рис. 11, *а*, или против часовой стрелки, рис. 11, *б*).

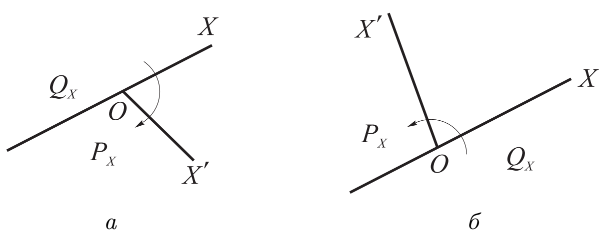


Рис. 11

Если точка O принадлежит F, то она сопоставляется сама себе. Точка O называется *центром* поворота.

Однако, как уже говорилось, в геометрии нет часовых стрелок, поэтому данное определение нужно дополнить определением того, что

значит «поворот в одну сторону». Это нужно определить для поворота на угол, меньший 180°.

Рассмотрим два отрезка OX, OY, которым сопоставлены отрезки OX', OY', образующие с ними неразвернутые углы. Ту полуплоскость, ограниченную прямой OX, где лежит отрезок OX', обозначим P_X; дополнительную полуплоскость обозначим Q_X. Тот же смысл для отрезка OY имеют обозначения P_Y, Q_Y.

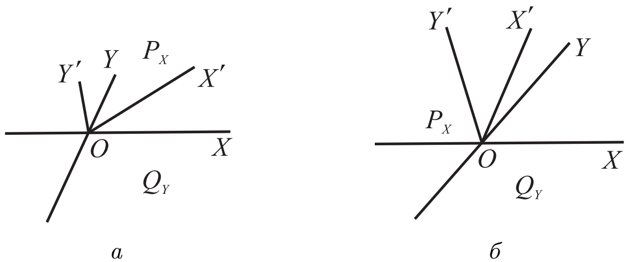


Рис. 12

Даем определение. Отрезки OX, OY, «переходя» в OX', OY', поворачиваются в *одну сторону*, когда выполнено следующее:

1) если $OY \subset P_X$, то $OX \subset Q_Y$ (рис. 12, *а, б*),

2) если же $OY \subset Q_X$, то $OX \subset P_Y$ (рис. 13, *а, б*). Первое означает: если отрезок OY лежит с той же стороны от OX, что и OX', то отрезки OY' и OX лежат с разных сторон от OY. (Наглядно это означает, что отрезок OY поворачивается в направлении от OX.) Во втором случае соотношение обратное.

Хотя в данном определении отрезки OX, OY внешне играют разную роль, оно на самом деле совершенно симметрично. При перемене X и Y местами получаем то же самое.

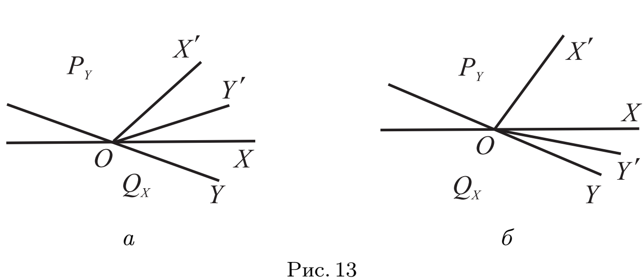


Рис. 13

(Например, поменяем ролями X и Y в условии 1), так что пусть $OX \subset P_Y$. Тогда не может быть $OY \subset P_X$, потому что в этом случае по 1) было бы $OX \subset Q_Y$. Стало быть, если $OX \subset P_Y$, то $OY \subset Q_X$.)

Определим еще повороты с помощью полярных координат.

Эти координаты задаются началом O, начальным лучом и направлением отсчета углов. А оно задается указанием полуплоскости, ограниченной лучом α (т. е. прямой, содержащей луч a); если сторона b угла ab лежит в этой полуплоскости, угол считается положительным, в противном случае он считается отрицательным (или сверхтупым).

Определяем поворот. Пусть O — центр поворота; примем его за начало полярных координат r, θ. Поворот фигуры вокруг O на угол α в том же направлении, в каком отсчитывается угловая координата, состоит в том, что каждой точке $X(r,\ \theta)$ фигуры F сопоставляется точка $X'(r,\ \theta + \alpha)$. Поворот в обратную сторону сопоставляет точке $X(r,\ \theta)$ точку $X'(r,\ \theta - \alpha)$. (Если углу α приписывать знак тот же, что и θ, то можно считать, что точке $X(r,\ \theta)$ сопоставляется $X'(r,\ \theta + \alpha)$; рис. 14.)

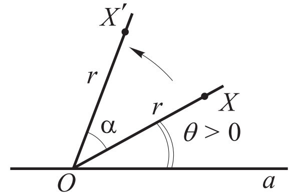


Рис. 14

Это определение не зависит от выбора начального луча системы координат r, θ. При замене начального луча a угловая координата θ заменяется на $\theta' = \theta + \theta_0$. Поэтому поворот, сопоставляя углу θ угол $\theta + \alpha$, сопоставляет углу $\theta' = \theta + \theta_0$ угол $(\theta + \alpha) + \theta_0 = \theta' + \alpha$.

Поворот. Таким образом, поворот состоит в том, что угловая координата изменяется на угол поворота. Отсюда очевидно следует, что композиция поворотов вокруг одного центра представляется сложением углов этих поворотов (конечно, с учетом знака и с точностью до слагаемых, кратных 2π). Обратный поворот получается при перемене знака угла.

Теорема 1. *Поворот вокруг точки сохраняет расстояния.*

Доказательство. Пусть при повороте вокруг точки O точки X, Y перешли в X', Y'. Введем полярные координаты с центром O. При повороте угловые координаты θ, θ_1 точек X, Y получают одно и то же слагаемое — угол поворота α. Поэтому если θ', θ_1' — координаты точек X', Y' то

$$\theta' - \theta_1' = \theta - \theta_1.$$

Эти разности — это углы между OX и OY, OX' и OY', они, стало

быть, равны. А так как $OX = OX'$, $OY' = OY$, то так же $X'Y' = XY$, что и требовалось доказать. □

Поворот в пространстве. Поворот в пространстве происходит вокруг прямой, называемой *осью* поворота, и, коротко говоря, состоит в том, что в плоскостях, перпендикулярных оси, происходит поворот на один и тот же угол в одну сторону вокруг точки пересечения плоскости с осью (рис. 15).

Точнее, *поворот* фигуры F *вокруг прямой* a — это отображение, при котором каждая точка $X \in F$ отображается в точку X' так, что 1) проекция точек X, X' на прямую a одна и та же; 2) если это точка O, то отрезки OX, OX' равны; 3) углы, образуемые отрезками OX, OX' равны для всех X; 4) все отрезки OX, переходя в OX', поворачиваются в одну сторону. Последнее означает, что проекции любых двух отрезков OX, O_1X_1 на плоскость, перпендикулярную прямой a, пово-

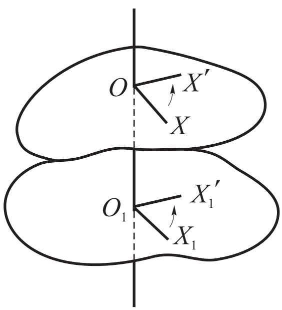


Рис. 15

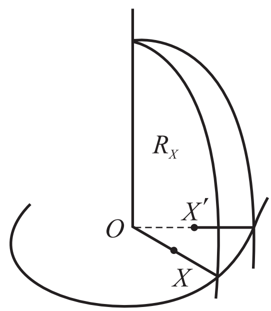


Рис. 16

рачиваются в одну сторону (здесь O_1 — проекция точки X_1 на a). Это можно представить еще так. Через каждую точку X, не лежащую на прямой a, проходит полуплоскость R_X, ограниченная этой прямой. Такая полуплоскость пересекает любую плоскость α, перпендикулярную a, по лучу с началом в точке O, в которой α пересекается с a. Поворот в плоскости α определяет поворот полуплоскостей R_X. И так во всех плоскостях α происходят согласованные повороты; о них говорится, что они происходят *в одном направлении* (рис. 16).

Теорема 2. *Поворот вокруг прямой сохраняет расстояния.*

Доказательство. Пусть при повороте вокруг прямой a точки A, B перешли в A', B'. Если они лежат в одной плоскости, перпендикулярной a, то $AB = A'B'$, поскольку в каждой такой плоскости происходит поворот.

Пусть точки A, B лежат в разных плоскостях α, β, перпендикулярных a, и переходят в точки A', B'. Пусть C — проекция точки A на плоскость β. Можно, конечно, считать, что точка C также «поворачивается» вместе с другими в плоскости β и переходит в C'. Тогда как $AC||a$, так и $A'C'||a$, так что AC и $A'C'$ перпендикулярны обеим плоскостям α, β. Стало быть, $A'C' = AC$ (так как $\alpha||\beta$, то отрезки AC, $A'C'$ равны как общие перпендикуляры параллельных плоскостей). Поэтому если C совпадает с B, то $A'B' = AB$.

Допустим, C не совпадает с B. Тогда $BC \perp AC$ (рис. 17) (так как $AC \perp \beta$). Треугольник ABC прямоугольный и

$$AB^2 = AC^2 + BC^2.$$

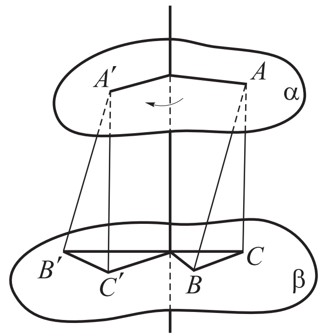


Рис. 17

Треугольник $A'B'C'$ тоже прямоугольный, так как $A'C' \perp \beta$. Кроме того, $A'C' = AC$ и $B'C' = BC$, поскольку в плоскости β происходит поворот. Из всего этого следует, что

$$A'B'^2 = A'C'^2 + B'C'^2 = AC^2 + BC^2 = AB^2.$$

Итак, $A'B' = AB$, что и требовалось доказать. □

О преобразованиях говорят, что они *коммутирующие* (*перестановочные*), если их композиция не зависит от порядка, в котором они производятся. Повороты вокруг одной оси коммутируют; их композиция соответствует сложению углов поворота (со знаком).

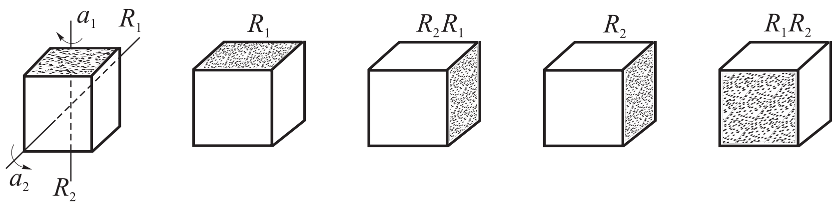


Рис. 18

Композиция двух поворотов вокруг пересекающихся осей, как будет доказано, представляет собою поворот вокруг третьей оси, проходящей через точку пересечения данных осей. Но повороты вокруг

разных осей, вообще говоря, не коммутируют. Если R_1, R_2 — повороты вокруг пересекающихся осей, то, вообще говоря, повороты R_2R_1 и R_1R_2 происходят вокруг разных осей. На рис. 18 изображены повороты куба вокруг осей a_1, a_2.

Винтовое наложение. *Винтовым наложением* называется композиция поворота и переноса вдоль оси поворота (т. е. на вектор, ей параллельный).

Порядок переноса и поворота при этом безразличны: они коммутируют.

Винтовое наложение, как композиция наложений, действительно является наложением.

Перенос и поворот можно рассматривать как частные случаи винтового наложения: с нулевым поворотом и с нулевым переносом.

Зеркальный поворот. *Зеркальным поворотом* называется композиция поворота и отражения относительно плоскости, перпендикулярной оси поворота.

Как композиция наложений зеркальный поворот является наложением.

Пусть a — ось поворота, α — плоскость, в которой происходит отражение, O — их точка пересечения. При отражении все точки плоскости α неподвижны, так что в ней происходит только поворот вокруг точки O. И он определяет поворот вокруг оси (как отмечено выше). Отсюда, между прочим, ясно, что порядок поворота и отражения в зеркальном повороте безразличен: они коммутируют.

Угол поворота, входящий в зеркальный поворот, называется *углом* этого *зеркального поворота.* Отражение можно рассматривать как зеркальный поворот с нулевым углом поворота.

Противоположный случай — *зеркальный поворот* с наибольшим возможным углом, т. е. *с поворотом на все* 180°, *представляет собою центральную симметрию с центром* O.

Действительно, возьмем какую угодно точку A, и пусть B, C — ее проекции на плоскость α и ось a, так что $\overrightarrow{OA} = \overrightarrow{OB} + \overrightarrow{OC}$. При повороте на 180° в плоскости α вектор $\overrightarrow{OB}$ переходит в $\overrightarrow{OB'} = -\overrightarrow{OB}$, а $\overrightarrow{OC}$ не изменяется. Но при отражении относительно плоскости α вектор $\overrightarrow{OB'}$ уже не изменяется, а вектор $\overrightarrow{OC}$ переходит в $\overrightarrow{OC'} = -\overrightarrow{OC}$. Поэтому в результате $\overrightarrow{OA}$ переходит в $\overrightarrow{OA'}$, так что

$$\overrightarrow{OA'} = \overrightarrow{OB'} + \overrightarrow{OC'} = -\overrightarrow{OB} - \overrightarrow{OC} = -\overrightarrow{OA}.$$

Это и означает, что происходит симметрия — отражение относительно центра O. □

Замечательно, что здесь расположение плоскости отражения и оси поворота не играет роли, лишь бы они были перпендикулярны и пересекались в точке O. При любом таком их расположении получается центральная симметрия.

Пример. Пример фигуры, совмещающейся сама с собой зеркальным поворотом, представляет многогранник, называемый «антипризмой». Он строится так.

Возьмем правильный n-угольник P_1 и проведем через его центр прямую a, перпендикулярную его плоскости α_1. Возьмем любую плоскость $\alpha || \alpha_1$ и произведем зеркальный поворот, состоящий из отражения относительно плоскости α и поворота вокруг прямой a на угол $\varphi = 2\pi/2n$ (в любую сторону). Такой зеркальный поворот f переводит многоугольник P_1 в многоугольник P_2, лежащий в плоскости $\alpha_2 || \alpha_1$ и имеющий центр в точке ее пересечения с осью a.

Вершины многоугольника P_2 будут расположены над лучами, идущими из O через середины сторон многоугольника P_1 (рис. 19). Соединяя вершины многоугольника P_2 с концами соответствующих сто-

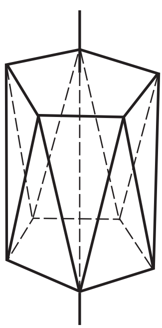


Рис. 19

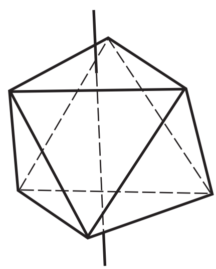


Рис. 20

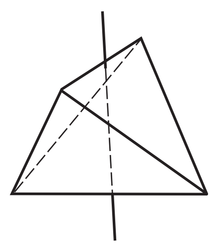


Рис. 21

рон P_1, получаем сеть ребер многогранника. Этот многогранник — «антипризма» — переводится сам в себя зеркальным поворотом f, в чем непосредственно убеждаемся из того, что его основания переводятся друг в друга этим зеркальным поворотом. Но этот многогранник совмещается сам с собою *поворотом* только на угол 2φ. Это, между прочим, служит основанием рассматривать зеркальный поворот как одно преобразование, не разлагая его на поворот и отражение.

Простейший случай «антипризмы» получаем, когда P_1 — треугольник; тут «антипризмой» является октаэдр (рис. 20).

Можно сказать, еще проще, когда P_1 — «двуугольник», т. е. отрезок, тогда получаем тетраэдр (рис. 21). Убедитесь!

Более наглядно называть антипризму «скрученная призма».

§ 3. Основные теоремы о наложениях. Их классификация и композиции

Формулировка теорем.

Теорема 1. *Всякое наложение в пространстве представляет собою одно из трех:*

а) *винтовое наложение, включая частные случаи переноса и поворота;*

б) *скользящее отражение, включая частный случай отражения;*

в) *зеркальный поворот, в частности центральную симметрию.*

Это значит, что любые две равные фигуры совмещаются одна с другой одним из указанных наложений. Или еще: каким наложениям ни подвергать фигуру, как ее ни двигать, ни крутить, ни отражать в плоскостях, а результатом будет одно из трех указанных наложений.

В планиметрии выполняется сходная

Теорема 2. *Всякое наложение в плоскости представляет собою одно из трех:*

а) *перенос*;

б) *поворот*;

в) *скользящее отражение, в частности одно отражение относительно прямой.*

Эта теорема является прямым следствием предыдущей, потому что всякое наложение в плоскости можно рассматривать как наложение в пространстве. Достаточно представить плоскость погруженной в пространство и из наложений, возможных в пространстве, выделить те, которые сохраняют плоскость — отображают ее саму на себя. □

Теоремы 1, 2 о конкретном виде наложений дополняются теоремами о композиции.

Теорема 3. *В пространстве всякая композиция переносов и поворотов (и, стало быть, вообще винтовых наложений) представляет собой винтовое наложение. Включение же в нее отражения дает либо зеркальный поворот, либо скользящее отражение.*

Теорема 4. *На плоскости композиция переносов и поворотов представляет собою либо перенос, либо поворот. Включение же в нее отражения дает скользящее отражение.*

Дальше мы уточним эти теоремы, указав в каких случаях получается то или иное наложение. Теоремы 3, 4 (с их упомянутыми уточнениями) оказываются следствиями следующей замечательной теоремы.

Теорема 5. *Всякое наложение является композицией отражений: в пространстве — относительно плоскостей, на плоскости — относительно прямых. Перенос и поворот (одинаково на плоскости и в пространстве) являются композициями двух отражений. Винтовое наложение — композицией четырех отражений. Скользящее отражение и зеркальный поворот — композицией трех отражений*[1].

Отсюда следует:

Композиция четного числа отражений представляет собой перенос, поворот или винтовое наложение; композиция нечетного числа отражений представляет собою скользящее отражение или зеркальный поворот. Отсюда легко получить и такое следствие.

Обозначим для краткости наложения перечисленных в теореме 1 трех типов цифрами I, II, III и их композиции обозначим как произведения I × I и т. п. Тогда можно утверждать: композиции I × I, II × II, III × III, II × III дают I, а композиции I × II, I × III дают II или III.

Наложения первого и второго рода. Наложения делятся на два класса: переносы, повороты и винтовые наложения называются наложениями *первого рода*; наложениями *второго рода* называются скользящие отражения и — в пространстве — еще зеркальные повороты.

Наложение первого рода можно осуществить непрерывным движением, т. е. пусть, например, дано винтовое наложение S, слагающееся из переноса вдоль оси a на вектор $\boldsymbol{c}$ и поворота вокруг оси a на угол α. Мы представляем себе винтовые наложения S_t, соответствующие значениям параметра t от 0 до 1, с той же осью a, переносами на $\boldsymbol{c}t$ и поворотами на углы αt (в ту же сторону). Когда t изменяется от 0 до 1, то мы получаем непрерывное винтовое движение, приводящее к данному наложению S.

Напротив, отражение ***нельзя*** осуществить непрерывным рядом наложений. Поэтому наложения второго рода нельзя осуществить непрерывным движением.

[1]На плоскости нет ни винтового наложения, ни зеркального поворота, так что к ней относится здесь только скользящее отображение.

Рассмотрим наложения в пространстве. При определении векторного и смешанного произведения векторов было введено понятие ориентации трехвекторника. Наглядно различаются правые и левые трехвекторники, и наглядно очевидно, что непрерывным движением нельзя превратить правый трехвекторник в левый или наоборот: левый — в правый. Но при отражении такое «превращение» как раз и происходит. Строго ориентация трехвекторника определяется по отношению к данной системе координат знаком смешанного произведения. При наложении абсолютная величина смешанного произведения, очевидно, не изменяется (так как она равна объему параллелепипеда, построенного на трехвекторнике). При непрерывном движении знак, очевидно, не может измениться, а при отражении он изменяется на обратный! В частности, если векторы $\boldsymbol{a}$, $\boldsymbol{b}$ расположены в плоскости, в которой происходит отражение, а третий вектор $\boldsymbol{c}$ ей перпендикулярен, то при отражении он заменяется на $-\boldsymbol{c}$ и произведение $(\boldsymbol{abc})$ меняет знак.

Таким образом, наложения первого и второго рода различаются тем, что первые сохраняют ориентацию, а вторые ее изменяют. Кроме того, первые могут осуществляться непрерывным рядом наложений — непрерывным движением, для вторых это невозможно. Наложения первого рода соответствуют реальным перемещениям тел, наложения второго рода соответствуют отражению в зеркале, когда правое меняется на левое.

На плоскости различие наложений первого и второго рода аналогично.

Доказательство основной теоремы о наложениях в пространстве. Здесь мы докажем теорему, из которой теорема 1 вытекает непосредственно благодаря теореме 3 о композициях. А эта теорема будет доказана в следующем параграфе.

Теорема 6. *Наложение однозначно задается наложением четырех точек, не лежащих в одной плоскости.* Подробнее:

Пусть в пространстве даны четыре точки A, B, C, D, не лежащие в одной плоскости, и еще такие точки A', B', C', D', что $A'B' = AB$, $B'C' = BC$ и т. д., так что отображение точек A на A', B на B' и т. д. представляет собою наложение. Тогда для любой фигуры, содержащей точки A, B, C, D, и, в частности, для всего пространства существует, и притом единственное, наложение, при котором происходит указанное наложение точек A, B, C, D.

Это наложение можно получить композицией переноса, двух поворотов и, может быть, еще отражения.

Доказательство. Пусть даны точки A, B, C, D, не лежащие

в одной плоскости, и такие точки A', B', C', D', что $A'B' = AB$ и т. д. Будем переводить точки A, $\ldots$, D в A', $\ldots$, D' наложениями всего пространства.

1) Произведем перенос, который переведет точку A в A'. Остальные точки перейдут в некоторые точки B_1, C_1, D_1, причем равенства отрезков сохранятся: $A'B_1 = AB$, $B_1C_1 = BC$ и т. д.

2) Теперь переведем B_1 в B' поворотом вокруг оси, проходящей через точку A'. Ось берется, понятно, перпендикулярной плоскости $A'B'B_1$ (рис. 22). (Если B_1 и без того совпадает с B', то поворот не нужен или, формально, — он тождественный.)

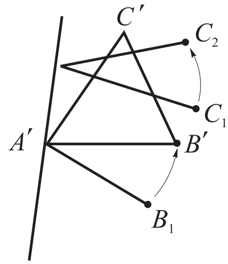


Рис. 22

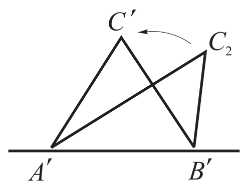


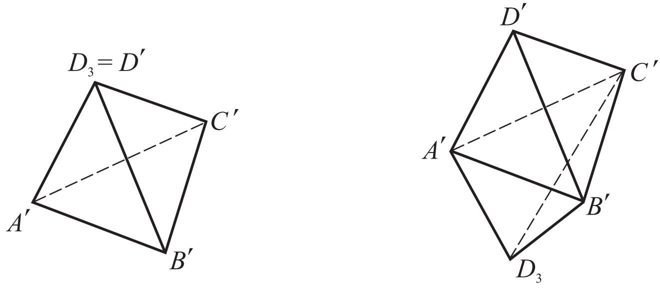
Рис. 23

При этом повороте точки C_1, D_1 перейдут в какие-то C_2, D_2. Равенства отрезков сохранятся.

3) Теперь переводим C_2 в C', оставляя A', B' на месте. Для этого производим поворот вокруг прямой $A'B'$, который переводит C_2 в C' (рис. 23). Это возможно, так как треугольник $A'B'C_2$ равен треугольнику $A'B'C'$ по трем сторонам.

Точка D_2 перейдет в некоторую точку D_3. Равенство отрезков сохранится, так что будет $D_3A' = D'A'$, $D_3B' = D'B'$, $D_3C' = D'C'$ (рис. 24). Поэтому для точки D_3 есть только две возможности: либо она совпадает с D' (рис. 24, *а*), либо расположена по другую сторону от плоскости $(A'B'C')$ и симметрична D' относительно этой плоскости (рис. 24, *б*). (Иначе говоря, если два тетраэдра $A'B'C'D'$ и $A'B'C'D_3$ с равными ребрами имеют общую грань, то они симметричны относительно плоскости $(A'B'C')$. Это утверждение читатель докажет сам.) Во втором случае точка D_3 переводится в D' отражением.

Итак, мы перевели точки A, B, C, D в точки A', B', C', D' переносом, двумя поворотами и, может быть, еще отражением. Так как все эти отображения представляют наложение пространства на себя



а *б*

Рис. 24

и композиция наложения есть опять наложение, то, значит, доказано, что точки A, B, C, D переводятся в A', B', C', D' наложением всего пространства.

Остается доказать, что такое наложение единственное. А это вытекает из следующего простого утверждения.

Лемма. *Положение точки в пространстве однозначно определяется ее расстояниями от четырех точек, не лежащих в одной плоскости. Значит, если точки A, B, C, D перешли в определенные точки A', B', C', D', то и все точки пространства заняли определенное положение — наложение определено однозначно.*

Доказательство леммы просто. Точки, расположенные на данных расстояниях от каждой из точек A, B, C, D лежат на сферах с центрами в этих точках. Две сферы с центрами A, B пересекаются по окружности. Третья сфера с центром C пересекает эту окружность в двух точках, симметричных относительно плоскости (ABC). Расстояние до четвертой точки выбирает из этих двух симметричных точек одну. □

Таким образом, теорема 6 полностью доказана. □

По теореме 3 о композициях, композиция переноса и двух поворотов есть винтовое движение, а с добавлением отражения — зеркальный поворот либо скользящее отражение. Таким образом, мы получаем теорему 1 о конкретном виде всякого наложения. □

Наложения в плоскости. Здесь мы докажем теорему о наложениях в плоскости, аналогичную теореме 6, Из нее теорема 2 о виде наложений в плоскости вытекает непосредственно благодаря теореме 4 о композициях. Саму же эту теорему мы докажем дальше в следующем параграфе.

Теорема 7. *Наложение плоской фигуры однозначно определяется наложением трех точек, не лежащих на одной прямой.* Подробнее:

Пусть на плоскости α даны три точки A, B, C, не лежащие на одной прямой, и еще такие точки A', B', C', что $A'B' = AB$, $B'C' = BC$, $C'A' = CA$, так что отображение точек A на A', B на B', C на C' представляет собою наложение. Тогда существует, и притом единственное, наложение всей плоскости α на себя, при котором происходит указанное наложение точек A, B, C.

Это наложение можно получить как композицию переноса, поворота и еще, может быть, отражения относительно прямой.

Доказательство. Пусть на плоскости даны точки A, B, C, не лежащие на одной прямой, и такие точки A', B', C', что $A'B' = AB$, $B'C' = BC$, $C'A' = CA$. Будем переводить точки A, B, C в A', B', C' наложениями всей плоскости.

1) Произведем перенос, которым переведем точку A в A'. Точки B, C перейдут при этом в какие-то точки B_1, C_1, причем равенство отрезков сохранится, так что $A'B_1 = AB$, $B_1C_1 = BC$, $A'C_1 = AC_1$.

2) Поворотом вокруг точки A' переведем точку B_1 в B'.

Точка C_1 перейдет в некоторую точку C_2, причем равенство отрезков сохранится, так что будет $A'C_2 = A'C_1 = A'C'$, $B'C_2 = B'C_1 = B'C'$. Ввиду этих равенств треугольники $A'B'C'$ и $A'B'C_2$ равны. Поэтому для точки C_2 возможны только два положения: либо она совпадает с C', либо симметрична C' относительно прямой $A'B'$. В последнем случае нужно еще произвести отражение относительно этой прямой.

Итак, доказано, что точки A, B, C переводятся в A', B', C' композицией переноса, поворота и, возможно, отражения. Единственность наложения плоскости, переводящего точки A, B, C в A', B', C', следует из того, что положение точки на плоскости определяется расстояниями от трех точек, не лежащих на одной прямой. Доказательство этого очевидно. (Ср. с предыдущей леммой.) Теорема 7 доказана. □

По теореме 4 о композиции наложений, на плоскости композиция переноса и поворота есть поворот или перенос, а с добавлением отражения композиция дает скользящее отражение. Таким образом, мы еще раз получаем теорему 2 о наложениях плоскости. □

Вариант доказательства теоремы 1. Пусть, как и раньше, точки A, B, C, D не лежат в одной плоскости и точки A', B', C', D' таковы, что $A'B' = AB$, $B'C' = BC$ и т. п. Покажем, как путем отражений перевести A, B, C, D в A', B', C', D'.

Произведем отражение r_A, переводящее A в A' (относительно какой

плоскости?). Точки B, C, D перейдут в некоторые B_1, C_1, D_1. (Если точки A и A' совпадают, то можно произвести отражение относительно любой проходящей через них плоскости. Так же действуем и далее.)

Произведем отражение r_B, переводящее B_1 в B. Оно происходит относительно плоскости, проходящей перпендикулярно $B'B_1$ через точку A' и биссектрису угла между $A'B'$ и $A'B_1$, поскольку $A'B' = A'B_1$. Точка A' останется на месте, B_1 перейдет в B', а C_1, D_1 перейдут в некоторые C_2, D_2.

Произведем отражение r_C, переводящее C_2 в C'. Оно происходит относительно плоскости, проходящей через отрезок $A'B'$, так что точки A' и B' остаются на месте (убедитесь!).

Итак, в результате трех отражений r_A, r_B, r_C точки A, B, C отобразились на A', B', C'. При этом для точки D есть две возможности: 1) точка D отобразилась на ту сторону от плоскости $(A'B'C')$, где лежит D', и тогда она совпала с D'; 2) точка D отобразилась в точку D_3, лежащую с другой стороны от плоскости $(A'B'C')$.

В первом случае мы получили нужное наложение. Как композиция трех отражений r_A, r_B, r_C, оно является зеркальным поворотом или скользящим отражением.

Во втором случае, чтобы перевести точки A, B, C, D окончательно в A', B', C', D', нужно произвести еще отражение r_D в плоскости $(A'B'C')$. В этом случае наложение представляется композицией четырех отражений и, следовательно, представляет собой винтовое движение (в частности, перенос или поворот). $\square$

§ 4. Теоремы о композиции

Композиции наложений на плоскости. Здесь мы докажем теорему 4 из § 3 сначала в несколько неполном виде.

Теорема 1. *Композиция переноса и поворота есть перенос или поворот, а с добавлением отражения — скользящее отражение* (в теореме 4 § 3 говорится о любых композициях переноса и поворотов; для них доказательство будет дано дополнительно).

Вместе с этим будет установлена часть теоремы 5, § 3 о композициях отражений, относящаяся к плоскости.

Начнем с представления переноса и поворота композициями отражений. Доказательство соответствующих утверждений не составляет особого труда и предлагается читателю в качестве интересных упражнений.

Лемма 1. *Композиция двух отражений относительно параллельных прямых представляет собою перенос в перпендикулярном им направлении.* Если отражение происходит сначала относительно прямой a, потом — относительно b и расстояние между a и b равно d, то перенос происходит в направлении от a к b перпендикулярно этим прямым на расстояние $2d$ (рис. 25)[2]. Отсюда следует:

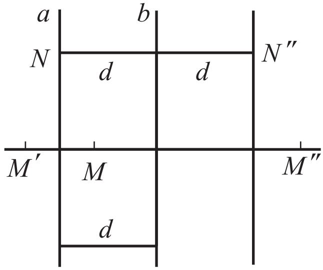


Рис. 25

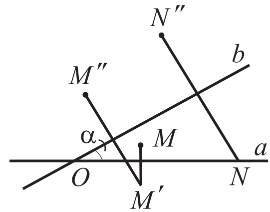


Рис. 26

а) Прямые, отражения относительно которых дают перенос, можно параллельно переносить (сохраняя расстояние между ними), а перенос будет получаться один и тот же.

б) Всякий перенос представи́м как композиция отражений относительно двух прямых, перпендикулярных направлению переноса.

Лемма 2. *Композиция отражений относительно двух пересекающихся прямых представляет собою поворот вокруг точки их пересечения.* Причем если отражение происходит сначала относительно прямой a, а потом — относительно прямой b и угол между ними равен α, то поворот происходит в направлении от a к b на угол 2α (рис. 26)[3]. Отсюда следует:

а) Всякий поворот представим как композиция отражений относительно двух пересекающихся прямых.

б) Прямые, отражения относительно которых дают данный поворот, можно поворачивать вокруг точки их пересечения, сохраняя угол между ними. Поворот будет получаться один и тот же. Поэтому одну из них всегда можно выбрать произвольно (лишь бы она проходила через центр поворота).

---

[2] Прежде всего заметим, что точки прямой a при первом отражении остаются на месте, а при втором как раз перемещаются на $2d$.

[3] Прежде всего заметим, что прямая a как раз поворачивается на 2α.

Доказательство теоремы 1. Докажем первую часть теоремы 1:

Композиция поворота (не тождественного) и переноса дает поворот.

Доказательство. Пусть сначала происходит поворот вокруг точки O, а потом перенос. Представим (согласно лемме 1) перенос отражениями сначала относительно прямой a, проходящей через O, а потом относительно другой прямой b. Поворот представим (согласно лемме 2) отражениями так, что второе отражение происходит относительно прямой a, первое — относительно некоторой прямой c. Тогда из этих отражений относительно c, a, a и b — останутся только отражения относительно c и b, т. е. поворот.

Если сначала происходит перенос — потом поворот, рассуждаем аналогично (проведите эти рассуждения).

Теперь докажем вторую часть теоремы 1:

Композиция переноса или поворота с отражением дает скользящее отражение (*в частности, одно отражение*).

Доказательство. Докажем это сначала для композиции переноса и отражения. Для этого разложим перенос в композицию двух: один из них производится вдоль прямой, относительно которой происходит отражение, другой — ей перпендикулярно. Этот последний вместе с отражением дает отражение относительно параллельной прямой (как непосредственно выводится из леммы 1).

Теперь докажем, что композиция поворота и отражения тоже дает скользящее отражение. Пусть сначала происходит поворот, а потом отражение относительно прямой a. Представим поворот композицией отражений относительно двух прямых, из которых вторая b параллельна a. Композиция отражений относительно прямых a и b дает перенос, так что мы получаем уже рассмотренную композицию переноса и отражения. При другом порядке: сначала отражение, потом поворот, вывод будет тот же самый.

Итак, теорема 1 доказана, а вместе с тем завершено доказательство теоремы 2, § 3 о видах отражений на плоскости. □

Композиции отражений в пространстве. Наложения, заданные на плоскости α, переносятся в пространство по перпендикулярам, и всякие наложения в пространстве, сохраняющие какую-либо плоскость и ограничиваемые ею полупространства, получаются таким путем. Отражению относительно прямой a соответствует при этом отражение относительно плоскости (перпендикулярной α и проходящей через прямую a). Переносу отвечает перенос, параллельный плоскости

α. Повороту — поворот вокруг оси, перпендикулярной плоскости α в центре поворота (рис. 27). Эти связи позволяют пересказать выводы предыдущего пункта применительно к пространству. Так, получаем

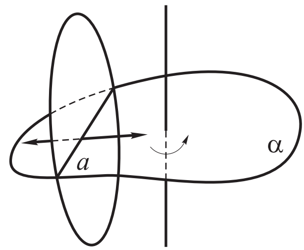


Рис. 27

1. *Композиция двух отражений относительно параллельных плоскостей представляет собой перенос в перпендикулярном им направлении.* Если отражение происходит сначала относительно плоскости α, потом — β, то перенос происходит от α к β (на удвоенное расстояние между α и β).

Всякий перенос можно представить как композицию отражений относительно двух параллельных плоскостей. Причем эти плоскости можно параллельно переносить, сохраняя расстояние между ними; отражения относительно них будут давать один и тот же перенос.

2. *Композиция отражений относительно двух пересекающихся плоскостей представляет собою поворот вокруг прямой их пересечения.* Всякий поворот представляется как композиция отражений относительно двух плоскостей; пересекающихся по оси поворота. Причем эти плоскости можно поворачивать вокруг оси, сохраняя угол между ними, и поворот будет один и тот же.

Для трех плоскостей есть две возможности расположения: они либо пересекаются в одной точке, либо параллельны одной прямой. Этому соответствуют два утверждения о композиции трех отражений.

3. *Композиция отражений относительно трех плоскостей, параллельных одной прямой, представляет собой скользящее отражение* (в частности, одно отражение). И всякое скользящее отражение можно представить как композицию трех отражений. Это сводится к соответствующему утверждению для наложений в плоскости, перпендикулярной трем данным плоскостям. Два отражения дают перенос или поворот, и потому данное утверждение равносильно тому, что в плоскости композиция поворота или переноса с отражением есть скользящее отражение. (Подробности остаются читателю.)

На этом прямые следствия утверждений, касающихся наложений в плоскости, кончаются.

4. *Композиция трех отражений относительно плоскостей, пе-*

ресекающихся в одной точке, представляет собой зеркальный поворот.

Доказательство. Пусть отражения последовательно происходят относительно плоскостей α, β, γ, пересекающихся в одной точке. Повернем плоскости α, β вокруг прямой их пересечения так, чтобы плоскость β перешла в $\beta' \perp \gamma$ (α перейдет в некую α'). Теперь повернем β' и γ вокруг прямой их пересечения так, чтобы γ перешла в $\gamma' \perp \alpha'$ (β' перейдет в β''). Мы получили $\alpha' \perp \gamma'$ и $\beta'' \perp \gamma'$. Тут уже видно, что композиция отражений относительно этих плоскостей есть зеркальный поворот (рис. 28).

Композиция наложений в пространстве. Здесь мы докажем теорему 3, § 3 о композициях, хотя и не в полном объеме, но именно так, как она используется в доказательстве теоремы 1, § 3 о видах наложений в пространстве. Там наложение было представлено переносом и двумя поворотами вокруг пересекающихся осей с возможным добавлением отражения. Вот мы и докажем:

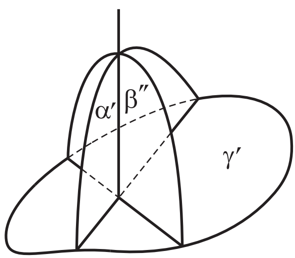


Рис. 28

Теорема 2. *Указанная композиция наложений без отражения, представляет собою винтовое наложение, а с добавлением отражения — скользящее отражение или зеркальный поворот.*

Доказательство разбивается на доказательства трех утверждений.

1. *Композиция двух поворотов вокруг пересекающихся осей представляет собой поворот* (*вокруг оси, проходящей через точку пересечения осей*).

Доказательство. Пусть два поворота происходят последовательно вокруг пересекающихся осей. Пусть α — плоскость, проходящая через эти оси. Тогда мы можем представить каждый из поворотов композицией двух отражений так, чтобы у каждого одно из отражений происходило относительно плоскости α. Первый поворот представляем так, чтобы это отражение было вторым, а второй — так, чтобы оно было первым. Первый поворот представляется отражениями сначала относительно подходящей плоскости β, а потом — α; второй — сначала относительно α, потом относительно подходящей плоскости γ. При последовательном осуществлении этих отражений отражения относительно плоскости α взаимно компенсируются. Поэтому останутся

только отражения относительно плоскостей β и γ. Они дают некоторый поворот, который, таким образом, и представляет композицию двух данных поворотов. □

2. *Композиция переноса и поворота представляет собою винтовое наложение.*

Доказательство. Пусть происходят последовательно перенос и затем поворот. Разложим перенос на два: один T_1 вдоль оси поворота, другой T_2 — перпендикулярно ей. В плоскости, перпендикулярной оси, перенос T_2 и поворот представляется (согласно связи наложений в плоскости и в пространстве) как перенос и поворот. По доказанному выше, их композиция представит поворот в плоскости, а он, в свою очередь, определяет в пространстве поворот вокруг оси, перпендикулярной плоскости и, тем самым, — параллельной оси исходного поворота.

Поэтому оставшийся перенос T_1, параллельный оси, дает в сочетании с этим поворотом винтовое наложение, что и требовалось. □

Если сначала производится поворот, потом — перенос, то вывод тот же.

3. *Композиция винтового наложения и отражения представляет собой зеркальный поворот или скользящее отражение.*

Доказательство. Пусть происходит композиция винтового наложения с отражением относительно плоскости α, пересекающей его ось в некоторой точке O. Представим перенос, входящий в данное винтовое наложение, отражениями относительно двух плоскостей β, β_1, из которых β_1 проходит через точку O. Поворот, входящий в данное винтовое наложение, представим как композицию отражений относительно двух плоскостей; обе они проходят через ось и значит — через точку O.

Таким образом, у нас получаются отражения относительно четырех плоскостей, проходящих через точку O, и еще отражение относительно плоскости β.

Соединяя попарно отражения относительно плоскостей, проходящих через точку O, получили два поворота, а их композиция представляет собой один поворот, т. е. два отражения.

Вместе с отражением в плоскости β это дает три отражения. А по утверждению 4 предыдущего пункта они дают зеркальный поворот или скользящее отражение. □

Таким образом, теорема 2 полностью доказана, а вместе с этим завершается доказательство теоремы 1, § 3 о видах наложений в пространстве. □

Завершение доказательства теорем о композициях наложений. Так как доказаны теоремы о видах наложений на плоскости и в пространстве, то тем самым доказано, что композиции наложений могут давать лишь наложения этих видов. А сведение наложений к композициям отражений вытекает из доказанного выше.

Без ссылки на теоремы о видах композиции можно рассудить так. Как доказано, в плоскости поворот с переносом дает поворот. Поэтому поворот можно представить как композицию поворота вокруг любого другого центра с переносом. В пространстве это позволяет переносить ось поворота, компенсируя это добавлением переноса, перпендикулярного оси поворота. Поэтому если даны два винтовых движения, то можно добавлением переноса привести к тому, что их оси будут пересекаться. Это дает один поворот в композиции с переносами. А это приводит к винтовому наложению.

Так любую композицию винтовых наложений, объединяя их попарно, можно привести к винтовому наложению.

Добавление отражения дает, как уже доказано, скользящее отражение или зеркальный поворот. □

§ 5. Симметрия

Общее понятие симметрии. *Симметрией* фигуры вообще называется свойство фигуры, состоящее в том, что существует ее (нетождественное) наложение самой на себя. Само слово «симметрия» происходит от греческого и означает в переводе соразмерность.

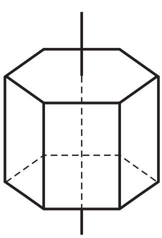


Рис. 29

Примеры симметрии плоских фигур дают правильные многоугольники. Примеры симметрии пространственных фигур дают правильные призмы и пирамиды: они совмещаются сами с собой, например, поворотами вокруг оси, перпендикулярной плоскости основания и проходящей через его центр (рис. 29).

Понятие симметрии понимают нередко в более узком смысле, включая в него на плоскости только симметрию относительно прямой и центральную симметрию, а в пространстве — симметрию относительно плоскости, относительно прямой и центральную симметрию. Соответствующие наложения сами себе обратны, и в фигуре с такой симметрией каждой точке A соответствует симметричная точка A', причем сама A симметрична A'. Симметрия относительно

прямой a в пространстве означает, что фигура совмещается сама с собой при повороте на 180° вокруг оси a. Каждой точке A сопоставляется при этом такая точка A', что отрезок AA' перпендикулярен оси и делится ею пополам.

Однако мы будем понимать симметрию в общем смысле, как она определена в начале и как ее понимают, в частности, когда говорят о симметрии кристаллов. При этом наложения фигуры на себя называются ***преобразованиями симметрии***.

Их основное свойство выражает

Теорема 1. *Композиция наложений фигуры на себя, так же как и отображение, обратное наложению фигуры на себя, является опять ее наложением на себя.*

Доказательство. Композиция наложений, как и отображение, обратное наложению, является наложением. Вместе с тем, если отображения S, S' отображают фигуру F на себя, то, произведя сначала отображение S, а за ним S', отобразим F на себя. Точно так же, если S отображает F на себя и обратимо, то обратное отображение, возвращая «все на прежние места», тем самым отобразит фигуру F на себя.

Из этих замечаний очевидно следует сказанное в теореме. □

Следствие. *Все наложения какой бы то ни было фигуры на себя образуют группу — ее* ***группу симметрии.***

Это следует из того, что совокупность отображений какого-либо множества на себя представляет собою группу тогда и только тогда, когда входящие в нее отображения обратимы, вместе со всяким отображением эта совокупность содержит ему обратное, а вместе с двумя отображениями — содержит их композицию (в любом порядке). Тождественное отображение, служащее единицей (нейтральным элементом) группы, принадлежит совокупности, как композиция любого из входящих в нее отображений с ему обратным. □

Взаимно однозначные или, что равносильно, обратимые отображения множества на себя называют также его ***преобразованиями***. Так, в частности, говорят о преобразованиях симметрии.

Замечание. Доказанная теорема и ее следствие имеют совершенно общее основание. Вообще, если взаимно однозначные отображения какого-либо множества на себя что-нибудь сохраняют, то их композиции и обратные отображения тоже это сохраняют (как, например, наложения сохраняют расстояния, вращения оставляют на месте центр или ось и т. п.).

Элементы симметрии ограниченных фигур. Фигура тем симметричней, чем обширнее ее группа симметрии. Самая симметричная фигура — все пространство: его группа симметрии — это группа всех вообще наложений. Но мы рассмотрим сначала симметрию ограниченных фигур.

Если фигура допускает какое-либо наложение R на себя, то она допускает (согласно теореме 1) и его композиции самого с собой: $R \cdot R$, $(R \cdot R) \cdot R$ и т. д. — произвольное число раз. Отсюда ясно, что наложения ограниченной фигуры на себя не могут включать ни переносов, ни винтовых наложений, ни скользящих отражений. Поэтому группа симметрии ограниченной фигуры может включать только повороты, отражения и зеркальные повороты; последние — только для пространственных фигур: на плоскости их нет.

Ось, вокруг которой происходят повороты, совмещающие фигуру саму с собой, называется ее ***осью симметрии***. Наибольшее число поворотов вокруг нее, совмещающих фигуру с собой, включая тождественный поворот, называется *порядком* оси. Так, у правильной n-угольной призмы есть ось симметрии n-го порядка, проходящая через центры оснований. Ось фигуры вращения — это ее ось симметрии бесконечного порядка.

Плоскость, отражение относительно которой совмещает фигуру саму с собой, называется ее ***плоскостью симметрии***. Ось поворота, входящего в зеркальный поворот, который совмещает фигуру саму с собой, называется ее ***зеркальной*** *осью симметрии. Порядком* этой оси также называется наибольшее число связанных с нею наложений — зеркальных и обычных поворотов, включая тождественный.

Все вместе эти оси и плоскости симметрии называют ***элементами симметрии*** *фигуры.* К ним присоединяется центр симметрии в качестве особого элемента симметрии (потому что, хотя центральная симметрия — это зеркальный поворот второго порядка, ось его неопределенная).

Зеркальный поворот вокруг данной оси a слагается из поворота вокруг оси a на некоторый угол φ и отражения относительно перпендикулярной ей плоскости α. Точка пересечения оси a и плоскости α служит при этом неподвижным центром. Повторение данного зеркального поворота приводит к повороту на угол 2φ, но без отражения в плоскости α, поскольку повторение отражения в данной плоскости дает тождественное преобразование.

В случае многогранника повторение зеркальных поворотов вокруг

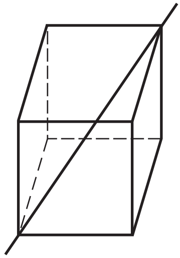


Рис. 30

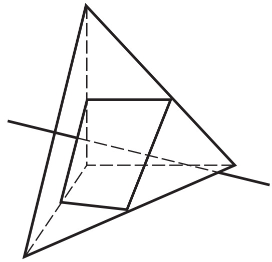


Рис. 31

зеркальной оси симметрии должно привести к тождественному преобразованию — к повороту на «все» 360°, без отражения. Поэтому число этих зеркальных поворотов четное. Зеркальная ось всегда четного порядка $2n$ и является вместе с тем осью симметрии n-го порядка с поворотами на 2φ.

У куба есть зеркальная ось 6-го порядка, проходящая по диагонали через противоположные вершины; она же ось 3-го порядка (рис. 30). У правильного тетраэдра есть зеркальная ось 4-го порядка, проходящая через середины противоположных ребер (рис. 31).

При наложении фигуры на себя ее элементы симметрии, очевидно, переходят в такие же элементы симметрии.

Все элементы симметрии ограниченной фигуры пересекаются в одной точке.

Если есть центр симметрии, то он и является этой точкой. Если бы какая-либо ось или плоскость симметрии не проходила через центр, то поворот вокруг такой оси или отражение в плоскости давали бы еще один центр. А композиция симметрии в двух центрах дает перенос. Фигура не могла бы быть ограниченной. Можно также заключить, что если бы, например, две оси не пересекались, то фигура не была бы ограниченной. (Доказательство мы оставляем читателям.)

Конечные группы симметрии — это группы симметрии правильных призм и правильных многогранников (которые мы рассмотрим в следующем параграфе), другие конечные группы симметрии — это подгруппы этих групп.

У фигур вращения имеется ось бесконечного порядка. Если же их две, то они пересекаются, и всякая прямая, проходящая через ту же точку, оказывается осью симметрии бесконечного порядка, как у шара.

Из выводов предыдущего параграфа о композиции поворотов вокруг пересекающихся осей следует, что композиции поворотов во-

круг двух осей симметрии порождают повороты вокруг третьей оси.

Симметрия неограниченных фигур. У неограниченных фигур могут быть элементы симметрии, связанные с переносами. Простейшие примеры представляют точки на прямой на равных расстояниях одна от другой, квадратная сетка на плоскости или в пространстве кубическая «решетка» (рис. 32). Фигура может также иметь винтовую симметрию, т. е. совмещаться сама с собой винтовым наложением (рис. 33), или симметрию скользящего отражения (рис. 34).

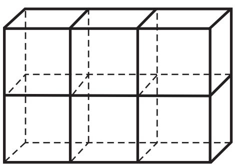


Рис. 32

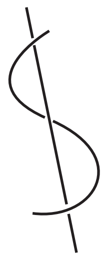


Рис. 33

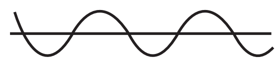
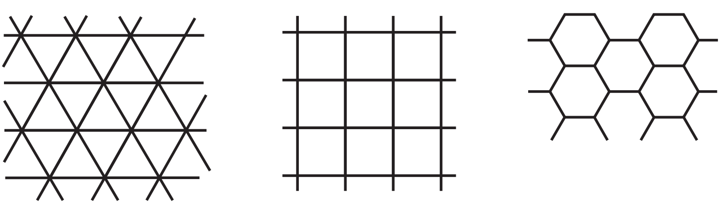


Рис. 34

Представляют особый интерес правильные системы фигур (в частности, точек) на плоскости и в пространстве — такие, в которых каждая фигура может быть совмещена с каждой наложением, совмещающим всю систему саму с собой. Особое значение правильных систем состоит в том, что они служат геометрическими моделями расположения атомов в кристаллах. В кристалле атомы или их комплексы образуют кристаллическую решетку, которая и изображается правильной системой фигур.

На плоскости примером правильной системы может служить «паркет» — система равных многоугольников, покрывающих всю плос-



а *б* *в*

Рис. 35

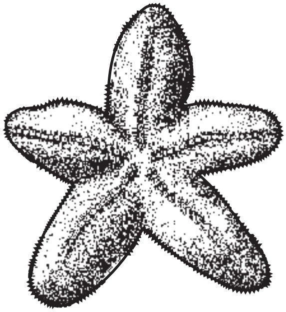


Рис. 36

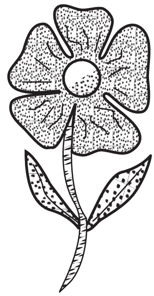


Рис. 37

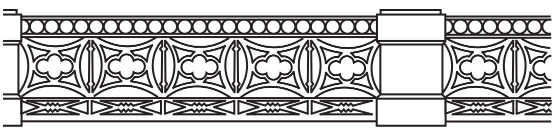


Рис. 38

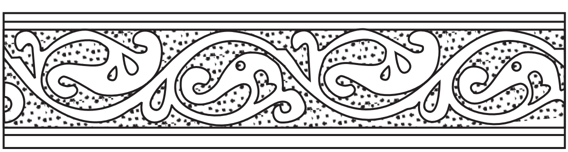


Рис. 39

кость, прилегая друг к другу по сторонам. Паркеты из правильных многоугольников составляются только из треугольников, квадратов и шестиугольников (рис. 35) (доказывается, например, из подсчета углов). Паркеты из равных выпуклых многоугольников с числом сторон больше 6 невозможны (можно также доказать из подсчета углов).

Симметрия в природе, в искусстве и т. д. Помимо кристаллов, симметрия в природе наблюдается у живых организмов. Подавляющее большинство животных, по крайней мере со стороны внешнего строения, имеют вертикальную плоскость симметрии; совсем особенную симметрию имеют морские звезды (рис. 36). У растений наблюдается симметрия цветов, симметрия листьев (рис. 37), в частности винтовая симметрия. Конечно, речь может идти лишь о конечном числе листьев; ограниченная фигура не может иметь, строго говоря, винтовую симметрию как самосовмещаемость винтовым движением. Но можно обобщить понятие симметрии, связанной с переносом, считая фигуру симметричной, если ее части допускают перенос или винтовое движение, совмещающее их с другими частями фигуры так, что при бесконечном продолжении фигура имела бы симметрию в исходном определении.

В таком же смысле можно говорить о переносной симметрии кристаллической решетки или чугунной решетки (рис. 38), бордюра или орнамента (рис. 39).

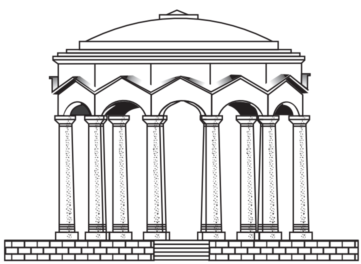


Рис. 40

В архитектуре симметрия зданий и фасадов играет существенную роль, как геометрическая основа прекрасного (рис. 40).

§ 6. Правильные многогранники

Виды правильных многогранников. Многогранник называется *правильным*, если у него все ребра равны, все плоские углы на гранях равны, все двугранные углы равны.

Из равенства ребер и плоских углов следует, что грани правильного многогранника представляют собой равные правильные многоугольники.

Аналогично из равенства плоских и двугранных углов следует, что все многогранные углы при вершинах равны. Многогранным углам соответствуют сферические многоугольники на единичной сфере. В данном случае у этих многоугольников стороны и углы равны. Поэтому они равны и являются правильными совершенно аналогично плоским многоугольникам. Они выпуклые, как и плоские правильные многоугольники.

Правильные многогранники определяют нередко несколько иначе (в школьных учебниках). Многогранник **правильный**, если он выпуклый, все его грани — равные друг другу правильные многоугольники и во всех вершинах сходится одинаковое число граней. Это определение равносильно предыдущему, но то определение ясно показывает, в каком смысле многогранник правильный: все его соответственные элементы — ребра и углы — равны (подобно правильному многоугольнику). Кроме того, при втором определении равенство всех двугранных углов требует особого доказательства[4] (о чем в школьных курсах не говорят).

Из школьного курса известно, что есть пять и только пять типов правильных многогранников:

1) тетраэдр (в переводе с греческого 4-гранник, рис. 41);
2) куб (гексаэдр — 6-гранник, рис. 42);
3) октаэдр (8-гранник, рис. 43);
4) додекаэдр (12-гранник, рис. 44);
5) икосаэдр (20-гранник, рис. 45).

Вопрос состоит в том, чтобы: (I) указать построение этих многогранников и тем самым убедиться, что такие правильные многогранни-

[4]Выполняется общая теорема: у выпуклых многогранников, одинаково составленных из соответственно равных граней, соответственные двугранные углы равны (теорема была доказана Коши и доказательство ее непросто). Можно сослаться на эту теорему. Это, очевидно, нужно только в случае, когда в вершинах сходится более 3-х граней.

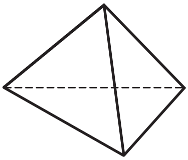


Рис. 41

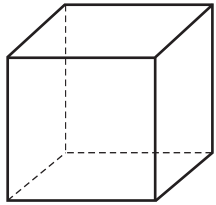


Рис. 42

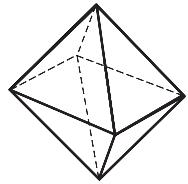


Рис. 43

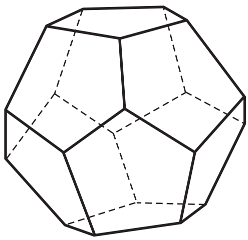


Рис. 44

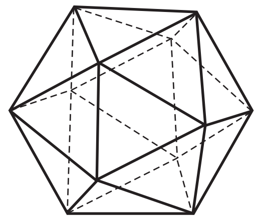


Рис. 45

ки действительно существуют; (II) доказать, что других правильных многогранников нет[5].

Рассмотрим оба вопроса совместно. Мы установили, что многогранные углы правильного многогранника выпуклы.

Сумма плоских углов выпуклого многогранного угла меньше 360°. Угол правильного n-угольника равен $180^\circ - 360^\circ/n$. Поэтому если выпуклый многогранный угол составлен из m таких углов, то должно быть

$$180m\left(1-\frac{2}{n}\right) < 360, \quad \text{т. е.} \quad m\left(1-\frac{2}{n}\right) < 2,$$

откуда

$$\frac{1}{m}+\frac{1}{n} > \frac{1}{2}.$$

Кроме того, очевидно $m \geqslant 3$ и $n \geqslant 3$. Поэтому m и $n < 6$, и есть только следующие возможности:

[5] Рассматривают еще «правильные многогранники» в другом смысле: такие, у которых «грани» — плоские замкнутые ломаные, — т. е. «звездчатые». Но это — многогранники в другом смысле.

n	3	3	3	4	5
m	3	4	5	3	3

Три случая треугольных граней, один — четырехугольных и один — пятиугольных.

Треугольники, сходящиеся в одной вершине, образуют боковую поверхность правильной пирамиды. Тетраэдр представляет собою правильную трехгранную пирамиду со всеми равными гранями.

Октаэдр составляется из двух пирамид с квадратным основанием. Другой возможности для того, чтобы в вершине сходилось по 4 треугольника, нет. (Если взять три квадрата, лежащие в трех взаимно перпендикулярных плоскостях, с попарно общими диагоналями, то их стороны дадут ребра октаэдра.)

Икосаэдр составляется из двух правильных пятиугольных пирамид, прилегающих основаниями к «скрученной призме» с пятиугольными основаниями и с 10 правильными треугольными боковыми гранями (рис. 46). (Впрочем, еще нужно убедиться, что все многогранные углы будут равны.)

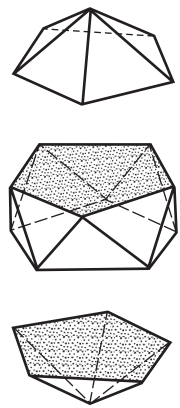


Рис. 46

Куб строится очевидным образом.

Для построения додекаэдра складываем два правильных пятиугольника по стороне так, чтобы к ним по двум сторонам можно было приложить еще один такой же пятиугольник. Этого, очевидно, можно добиться, если, сложив два пятиугольника, увеличивать затем двугранный угол между ними. По симметрии пятиугольник так же можно будет приложить к двум другим сторонам (у другого конца общей стороны; рис. 47). Так складывая пятиугольники, построим их пояс вокруг одного из них. Это даст половину додекаэдра, другая построится так же. Углы будут сходиться по доказанному — из симметрии.

Между правильными многогранниками есть двойственность. Центры граней одного служат вершинами другого, и обратно.

Тетраэдр двойственен тетраэдру (рис. 48), октаэдр — кубу и куб — октаэдру (рис. 49), икосаэдр — додекаэдру и додекаэдр — икосаэдру (рис. 50). Поэтому, построив додекаэдр, можно по центрам его граней построить икосаэдр.

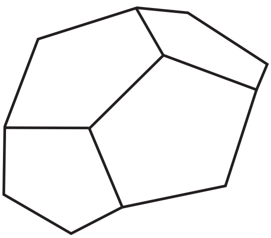


Рис. 47

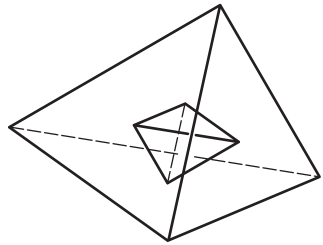
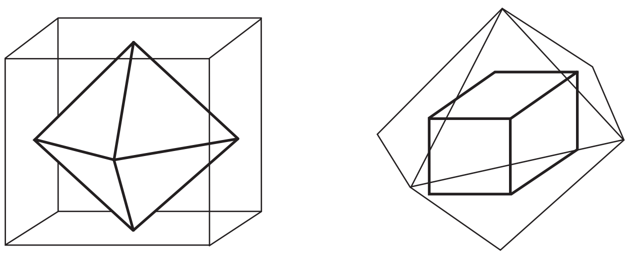
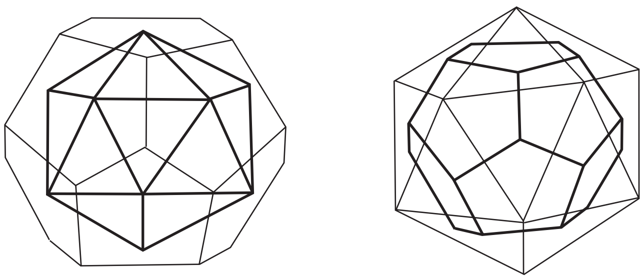


Рис. 48



а *б*

Рис. 49



а *б*

Рис. 50

Из построения додекаэдра ясно, что другого выпуклого, и тем более правильного, многогранника, составленного из правильных пятиугольников, быть не может. Многогранник, у которого по пять треугольников сходится в вершине, двойственен многограннику с пятиугольными гранями, т. е. додекаэдру. Следовательно, он — икосаэдр.

Симметрия правильных многогранников.

Теорема 1. *Рассмотрим данный правильный многогранник P. Пусть A — его вершина, a — ребро с концом A, α — грань со стороной a. Для любых других аналогичных его элементов A', a', α' существует наложение многогранника P на себя, переводящее A' в A, a' в a, α' в α.*

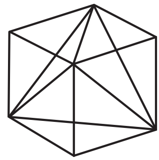
Доказательство. Переносом многогранника переведем вершину A' в A. Поворотом многогранника вокруг A переведем перенесенное ребро a' в a. Поворотом многогранника вокруг ребра a приведем (перенесенную и повернутую) грань α' в совпадение с гранью α. Так как грани равны, то грань α' полностью совместится с α.

Так как двугранные углы равны, то для граней β и β', смежных с α и α', есть только две возможности: 1) β' совпадает с β; 2) β' не совпадает с β, но симметрична β относительно плоскости грани α. В таком случае отражением в этой плоскости переведем β' в β.

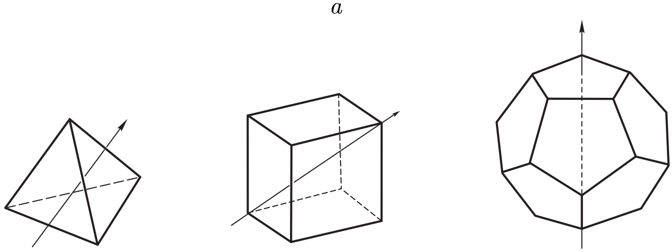
Итак, наложением всего многогранника P мы совместили вершину A' с A, ребро a' — с a, грани α', β', смежные по ребру a', — с гранями α, β, смежными по ребру a.

Убедимся, что при этом многогранник оказывается совмещенным сам с собой. Две грани многогранного угла при вершине A совпали (α' с α, β' с β). Перейдем к граням γ и γ', соседним с β. Двугранные углы, которые они образуют с β, равны и расположены с одной стороны — с той же, с какой лежит грань α. Поэтому грань γ' совпадает с γ. Так убедимся, что многогранные углы при вершине A совпали. Переходя к другой вершине, соединенной с A ребром, аналогично убедимся, что и при этой вершине многогранные углы совпадают. И так пройдя по всему многограннику, убедимся, что он совпал сам с собой, что и требовалось доказать. □

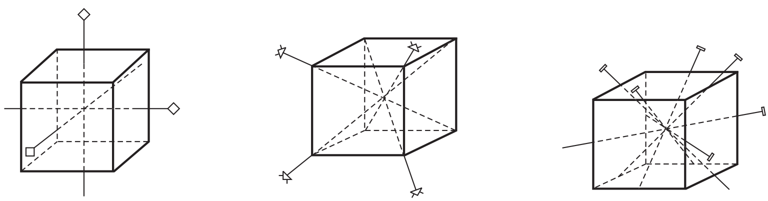
Свойство правильных многогранников, установленное доказанной теоремой, означает, что они обладают, так сказать, ***максимальной мыслимой симметрией***. Наложение, совмещение многогранника самого с собою, неизбежно совмещает какую-то вершину A' с A, ребро a' — с a, грань α' — с α и примыкающую грань β' — с β. Наложение этим вполне определено, оно только одно. Поэтому максимальное число возможных наложений будет тогда, когда каждую совокупность A,



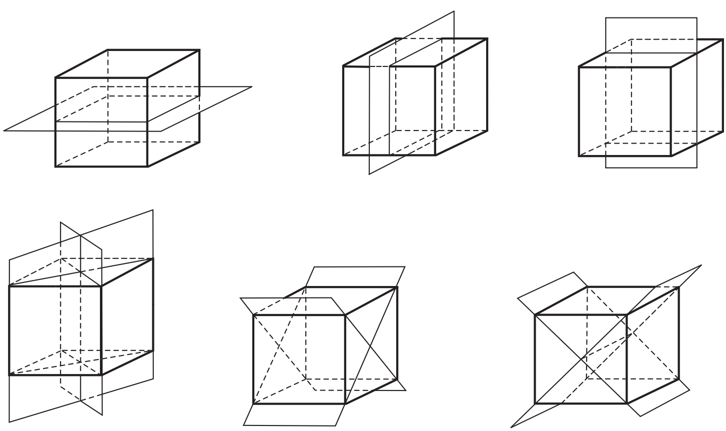
а



б



в



г

Рис. 51

a, α, β можно перевести в каждую. А это так у правильных многогранников.

Очевидно, верно и обратное. Если многогранник обладает такой максимальной симметрией, то он правильный (так как ребро a совмещается с a', угол на грани α' при вершине A совмещается с таким же углом, и двугранный угол между α' и β' совмещается с углом между α и β — так что все ребра и углы равны).

Число наложений, совмещающих правильный многогранник сам с собою, равно $2\,me$, где m — число ребер, сходящихся в одной вершине, и e — число вершин; me наложений первого рода и me — наложений второго рода. Они и образуют группу симметрии правильного многогранника.

Группы симметрии у куба и октаэдра совпадают ввиду их двойственности. Так же совпадают группы симметрии у додекаэдра и икосаэдра. Группа тетраэдра является подгруппой группы куба, как видно из возможности вложить тетраэдр в куб (рис. 51, а).

Наиболее интересные элементы симметрии — это зеркальные оси: 4-го порядка у тетраэдра, 6-го порядка — у куба, 10-го порядка — у додекаэдра (рис. 51, *б*). Убедитесь, что это так, определив, как расположены эти оси.

Оси симметрии и плоскости симметрии куба изображены на рис. 51, *в, г*. Рассмотрите элементы симметрии других правильных многогранников.

Глава II

ПОДОБИЯ И ИНВЕРСИИ

§ 1. Преобразования подобия

Общие свойства подобий. Отображение фигуры называется *отображением подобия* (или *подобием*), если при нем отношения расстояний между точками сохраняются, так что все расстояния изменяются в одно и то же число раз; выраженные в каком-либо масштабе, они умножаются на одно и то же число — *коэффициент подобия*.

Всякое наложение — это подобие с коэффициентом единица. Простейшее подобие, отличное от наложения, представляет гомотетия.

Гомотетия с центром O и коэффициентом $k \neq 0$ — это такое отображение, при котором всякая точка M переходит в такую M', что $\overrightarrow{OM'} = k \cdot \overrightarrow{OM}$ (рис. 52, *а, б* для $k > 0$, $k < 0$). В частности, при $k = -1$ гомотетия представляет собой центральную симметрию.

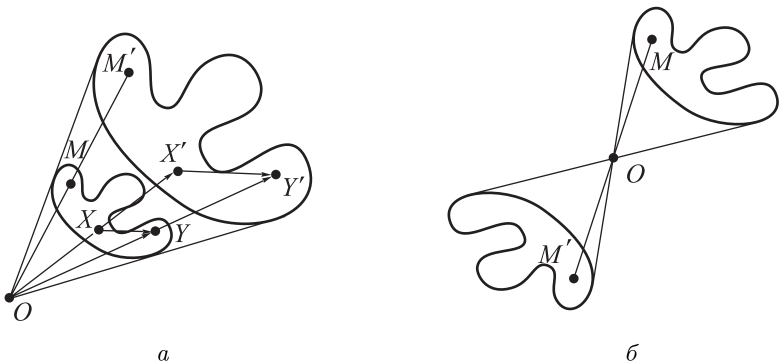


Рис. 52

При гомотетии с коэффициентом k все векторы умножаются на k.

Действительно, если при гомотетии с центром O и коэффициентом k точки X, Y перешли в X', Y', то (рис. 52)

$$\overrightarrow{OX'} = k\overrightarrow{OX}, \quad \overrightarrow{OY'} = k\overrightarrow{OY}.$$

А так как $\overrightarrow{XY} = \overrightarrow{OY} - \overrightarrow{OX}$ и $\overrightarrow{X'Y'} = \overrightarrow{OY'} - \overrightarrow{OX'}$, то

$$\overrightarrow{X'Y'} = k\overrightarrow{XY}. \quad \square$$

Поэтому гомотетия с коэффициентом k представляет собой подобное преобразование с коэффициентом $|k|$.

Теорема 1. *Всякое подобие является композицией гомотетии с тем же коэффициентом и наложения.*

Доказательство. Пусть фигура F путем подобия отображена на F' с коэффициентом подобия k. Подвергнем фигуру F гомотетии с тем же коэффициентом и с любым центром. Она перейдет в фигуру F''; и так как все расстояния умножаются на k, то они станут такими же, как в фигуре F'. Поэтому фигуру F'' можно отобразить на F' наложением. Таким образом, F отображается на F' в результате гомотетии и наложения, что и требовалось доказать. $\square$

Теорема 2. *Композиция подобий с коэффициентами k_1, k_2 есть подобие с коэффициентом k_1k_2. Отображение подобия обратимо, причем отображение, обратное подобию с коэффициентом k, является подобием с коэффициентом $1/k$.*

Доказательство очевидно. □

Теорема 3. *При подобном отображении отрезки отображаются на отрезки, причем равенство отрезков сохраняется, углы отображаются на равные, плоскости отображаются на плоскости; всё — как при наложениях* (см. § I.1).

Доказательство совершенно аналогично доказательству аналогичных свойств наложений (§ 1). □

Поскольку при преобразовании подобия плоскости, как и пространства, отрезки переходят в отрезки и равенство отрезков сохраняется, а в пространстве плоскости также переходят в плоскости, то все отношения, определяемые аксиомами, сохраняются. При этом подобные преобразования как плоскости, так и пространства, в силу теоремы 2, образуют группу. Поэтому можно сказать, что все свойства фигур евклидовой геометрии инвариантны (неизменны) относительно группы подобий.

Отметим в дополнение интересное свойство подобий.

Всякое подобное преобразование плоскости, как и пространства, отличное от наложения, имеет одну неподвижную точку. Оно представимо поэтому как композиция гомотетии с поворотом (может быть, тождественным), или с зеркальным поворотом вокруг оси, проходящей через центр гомотетии, или с отражением в плоскости, проходящей через центр гомотетии (впрочем, отражение можно считать частным случаем зеркального поворота, когда сам поворот тождественный).

§ 2. Инверсии

Определение. *Инверсией* (плоскости) относительно окружности с центром O и радиусом R называется такое отображение, при котором каждая точка M, отличная от O, отображается в такую точку M', что а) M' лежит на луче OM, б) $OM' \cdot OM = R^2$ (рис. 53).

Поскольку здесь

$$OM' = \frac{R^2}{OM}, \tag{1}$$

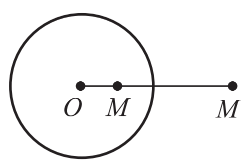


Рис. 53

то инверсию называют также *преобразованием обратных радиусов*.

Из (1) ясно, что если $OM = R$, то $OM' = OM$, и так как точка M' лежит на том же луче, идущем из O, что и точка M, то в данном

случае M' просто совпадает с M. То есть все точки окружности K с центром O и радиусом R остаются на месте.

Если же $OM < R$, то $OM' > R$, и если $OM > R$, то $OM' < R$. То есть внутренность круга с окружностью K отображается на его внешность, а внешность — на внутренность. Причем точки остаются на тех же лучах, идущих из O. Поэтому инверсия представляет собой как бы отражение в окружности. Ее так и можно назвать: отражение в окружности.

Точка O — центр инверсии — никуда не отображается. Когда точка M приближается к O: $OM \to 0$, то $OM' \to \infty$. Поэтому естественно считать, что центр O отображается в бесконечность. Дополняя плоскость бесконечно удаленной точкой, можно сказать: при инверсии центр отображается в бесконечно удаленную точку, а бесконечно удаленная точка — в центре[6].

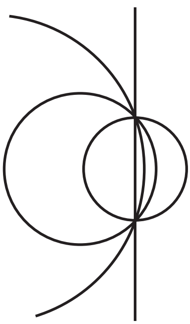


Рис. 54

Дополнив плоскость бесконечно удаленной точкой, будем считать, что прямая — это «окружность, проходящая через бесконечно удаленную точку». Пока мы условливаемся об этом чисто формально; хотя наглядно очевидно, что при бесконечном увеличении радиуса окружность, проходящая через две данные точки, «распрямляется» и в пределе переходит в прямую (рис. 54).

Приняв введенное условие о прямых, мы объединяем в понятие ***обобщенная окружность*** настоящие окружности и прямые. Это позволяет сформулировать важное свойство инверсии очень коротко.

Теорема 1. *Инверсия переводит обобщенные окружности в обобщенные окружности.*

В развернутом виде это означает, что инверсия переводит:

а) всякую настоящую окружность, не проходящую через центр инверсии, — в окружность такого же рода;

б) всякую настоящую окружность, проходящую через центр инверсии, — в прямую;

в) всякую прямую, не проходящую через центр, — в окружность, проходящую через центр;

г) всякую прямую, проходящую через центр,— саму в себя.

При этом в случаях б), в), г) нужно иметь в виду то условие, что

[6] Так же вводят бесконечно удаленную точку на комплексной плоскости в теории функций комплексной переменной.

центр переходит в бесконечно удаленную точку, а она — в центр; без этого нужно оговаривать, что центр не имеет ни образа, ни прообраза, так что он должен быть исключен.

Доказательство. Пусть $\boldsymbol{r} = \overrightarrow{OM}$, $\boldsymbol{r}' = \overrightarrow{OM'}$ радиус-векторы точки M и ее образа при инверсии в окружности K с центром O и радиусом a. Из определения инверсии

$$\boldsymbol{r} = a^2 \frac{\boldsymbol{r}'}{\boldsymbol{r}'^2}.$$

В прямоугольных координатах с началом O получаем

$$x = \frac{a^2 x'}{x'^2 + y'^2}, \quad y = \frac{a^2 y'}{x'^2 + y'^2},$$
$$x^2 + y^2 = \frac{a^2}{x'^2 + y'^2}. \tag{2}$$

Рассмотрим кривую, представляемую уравнением

$$b(x^2 + y^2) + cx + dy + e = 0. \tag{3}$$

Исключая случаи, когда это точка или пустое множество, кривая с таким уравнением — окружность, если $b \neq 0$, или прямая, если $b = 0$. Подставляя x и y из (2) и избавляясь от знаменателя, получим

$$a^4 b + a^2 c x' + a^2 d y' + e(x'^2 + y'^2) = 0, \tag{4}$$

т. е. уравнение того же вида, и, стало быть, представляющее окружность, если $e \neq 0$, и прямую, если $e = 0$.

Теорема доказана, и остается только просмотреть, когда получается каждый из четырех случаев, указанных в развернутой формулировке. Читатель сам легко это сделает. □

Инверсия сохраняет углы. Подробно понятие кривой и угла между кривыми обсуждается в части IV.

Лемма. *При инверсии, сопоставляющей друг другу точки A, A', окружность, проходящая через эти точки, отображается на себя.*

Доказательство. Если инверсия относительно некоторой окружности C сопоставляет друг другу точки A, A', то, значит, она проходит между ними, так что одна из точек A, A' находится внутри круга с окружностью C, а другая — вне. Поэтому если окружность K проходит через точки A, A', то она пересекает окружность C. Точка

пересечения B, как принадлежащая C, переходит сама в себя. Следовательно, окружность K переходит в окружность, проходящую через те же точки A, A', B. Но через три точки может проходить только одна окружность. Следовательно, окружность K отображается сама на себя, что и требовалось доказать. □

Теорема 2. *При инверсии углы между кривыми сохраняются, т. е. две кривые, пересекающиеся под некоторым углом, переходят при инверсии в кривые, пересекающиеся в соответственной точке под таким же углом.*

Доказательство. Пусть кривые L_1, L_2 пересекаются в некоторой точке A и при инверсии переходят в кривые L_1', L_2', пересекающиеся в соответствующей точке A'. Построим окружности K_1, K_2, проходящие через точки A, A' и имеющие в точке A те же касательные, что кривые L_1, L_2. Согласно доказанной лемме эти окружности перейдут при данной инверсии в себя. А угол, который они образуют друг с другом в точке A', равен углу, какой они образуют в точке A. Стало быть, и кривые, которых они касаются, образуют те же углы. □

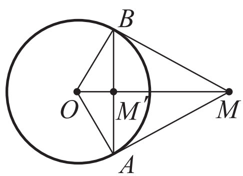


Рис. 55

Следующая теорема устанавливает такое свойство инверсии, которое позволяет строить для всякой данной точки ее образ при инверсии относительно данной окружности.

Теорема 3. *При инверсии относительно окружности круга K точки M, M', которые переходят друг в друга, связаны следующим образом. Если точка M' лежит внутри круга и служит серединой хорды AB, то M лежит вне круга, и прямые MA, MB служат его касательными* (рис. 55).

Доказательство. Пусть O — центр круга K. Треугольник OAM прямоугольный; OA — его катет, и если M' — середина хорды AB, то $AM' \perp OM$. Поэтому, согласно известному свойству высоты прямоугольного треугольника, $OM \cdot OM' = OA^2$, что и доказывает утверждение теоремы. □

Инверсия в пространстве. В пространстве инверсия определяется буквально так же, как на плоскости: задается точка O — центр инверсии и длина R — радиус инверсии, и каждой точке M ставится в соответствие точка M', лежащая на том же луче с началом O и такая, что $OM' \cdot OM = R^2$.

Все точки сферы радиуса R с центром O остаются на месте. Каждая плоскость, проходящая через O, отображается на себя, и в ней происходит инверсия с центром O и радиусом R.

Если дополнить пространство бесконечно удаленной точкой и считать плоскость «обобщенной сферой», проходящей через эту точку, то можно высказать теорему:

Теорема 4. *При инверсии «обобщенные сферы» переходят в «обобщенные сферы».*

Эта теорема подробно раскрывается подобно теореме 1. Доказательство ее аналогично. □

Из теоремы 4 следует, что и в пространстве «обобщенные окружности» переходят в «обобщенные окружности».

Теорема 5. *Инверсия сохраняет углы между кривыми.*

Доказательство аналогично доказательству теоремы 2. □

Замечание 1. Пусть M, M' — точки, соответствующие друг другу при инверсии относительно окружности C радиуса R, и пусть h, h' — их расстояния от C со знаком (считаем $h > 0$ для точки внутри окружности C и $h' < 0$ для точки вне окружности). Тогда выполняется равенство

$$\frac{1}{h} + \frac{1}{h'} = \frac{1}{R}.$$

Когда окружность C, «расширяясь», переходит в пределе в прямую c при $R \to \infty$, тогда $h' \to -h$, т. е. в пределе инверсия дает отражение относительно прямой c.

Замечание 2. Отображение, сохраняющее углы между кривыми, называется *конформным.* На плоскости существует чрезвычайно много конформных отображений. Так, внутренность круга допускает конформное отображение на любую односвязную область (т. е. такую, в которой всякая замкнутая кривая может быть стянута в точку). Это теорема Римана. Конформные отображения на плоскости представляются аналитическими функциями комплексной переменной, и всякая такая функция задает конформное отображение.

Но в пространстве всякое конформное отображение любой области является композицией инверсий и отражений относительно плоскостей — «инверсий» относительно сфер бесконечного радиуса. (Это теорема французского математика Ж. Лиувилля (1809–1882).)

Глава III

АФФИННЫЕ ПРЕОБРАЗОВАНИЯ И АФФИННАЯ ГЕОМЕТРИЯ

§ 1. Параллельное проектирование

Будем считать всякую прямую параллельной самой себе[7].

Аналогично, будем считать отрезки, лежащие на одной прямой, параллельными.

Определение. *Параллельным проектированием на плоскость* α *вдоль прямой* a, пересекающей α, называется отображение, сопоставляющее точке X ту точку X', в которой проходящая через X прямая, параллельная a, пересекает плоскость α (рис. 56). О прямой a говорят, что она задает *направление проектирования.*

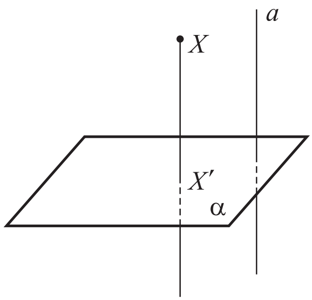


Рис. 56

Основные свойства параллельного проектирования выражаются следующей теоремой, известной еще из школьного курса.

Теорема 1. *При параллельном проектировании для прямых и отрезков, не параллельных направлению проектирования, выполняются следующие свойства.*

1. *Проекция прямой есть прямая, проекция отрезка — отрезок.*

2. *Проекции параллельных прямых параллельны.*

3. *Отношение длин проекций параллельных отрезков равно отношению длин самих этих отрезков. Другими словами, при параллельном проектировании отношения параллельных отрезков сохраняются.*

Мы докажем эту теорему, рассмотрев, что происходит при проектировании с векторами.

Отображения векторов. При отображении какой-либо фигуры F точкам сопоставляются точки, и соответственно фигурам, содержа-

[7]Это необходимо, когда рассматривается какое-либо множество параллельных прямых a, b, c и т. д. Известно, что если $a||b$ и $b||c$, то $a||c$. Поэтому если $a||b$, $b||c$, то $c||a$, и мы логически приходим к тому, что $a||a$. Когда мы говорим о прямых, параллельных данной прямой a, то естественно включать в них ее саму и говорить, что все они друг другу параллельны.

щимся в F, сопоставляются фигуры. А что происходит при этом с векторами — направленными отрезками? Даем определение.

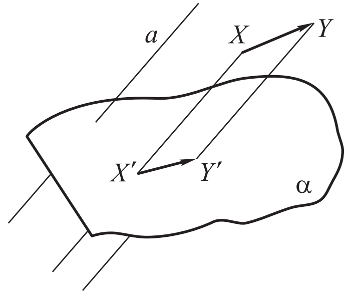


Рис. 57

Если при некотором отображении точки X, Y отображаются на X', Y', то вектору $\overrightarrow{XY}$ сопоставляется вектор $\overrightarrow{X'Y'}$. То есть образом данного вектора считается вектор, начало и конец которого являются образами начала и конца данного вектора $\overrightarrow{XY}$ (рис. 57).

(Что происходит при этом с точками, лежащими на отрезке XY, не играет роли; с ними может ничего не происходить, если отображение на них не распространяется, как, например, при отображении (сжатии) окружности в эллипс и т. п.)

Замечание. Вводя понятие вектора, мы различали конкретные и абстрактные — свободные — векторы. Конкретный вектор изображается направленным отрезком, но фактически в понятии о нем играют роль только его начало и конец. «Конкретный вектор» — это упорядоченная пара точек, а не фигура, и отрезок с нею связывается лишь для наглядности.

Для конкретных векторов определяется равенство (равные векторы представляют один абстрактный, свободный, вектор); операции с векторами определяются на основе перенесения вектора к любому началу; т. е. всякий данный конкретный вектор можно заменить равным вектором. Поэтому рассматривать такие отображения векторов, при которых нарушается их равенство, имеет мало смысла.

Векторы при параллельном проектировании. Заметим, что если мы хотим, пользуясь векторами, доказать, что при параллельном проектировании отрезки проектируются в отрезки, то рассматривать векторы как направленные отрезки нецелесообразно (раз еще нужно доказать, что отрезок отображается на отрезок). Еще и поэтому векторы нужно рассматривать как упорядоченные пары точек и понимать отображения векторов, как было только что определено.

Итак, рассмотрим проектирование вдоль прямой a на плоскость α.

Как было доказано, каждый вектор однозначно разлагается на составляющие: одну — параллельную плоскости α, другую — параллельную прямой a.

Теорема 2. *При разложении вектора $\overrightarrow{XY}$ на составляющие, параллельные плоскости α и прямой a, составляющие на плоскости α получаются параллельным проектированием вдоль прямой a. То есть если X', Y' — проекции точек X, Y, то вектор $\overrightarrow{X'Y'}$ и есть составляющая вектора $\overrightarrow{XY}$ на плоскости α* (рис. 57).

Доказательство. Для доказательства достаточно вспомнить, как строятся составляющие вектора, параллельные данной плоскости и данной прямой.

Пусть дан вектор — направленный отрезок $\overrightarrow{XY}$, — строим его составляющие вдоль прямой a и плоскости α.

Через начало вектора $\overrightarrow{XY}$ проводим плоскость β, параллельную α, и из конца вектора проводим прямую b, параллельную прямой a. Точка Y'' ее пересечения с плоскостью β и дает конец вектора $\overrightarrow{XY''}$, служащего составляющей вектора $\overrightarrow{XY}$ на плоскости β. Проводя через X и Y прямые, параллельные a, до пересечения с плоскостью α, получаем в ней точки X', Y'. При этом, как очевидно[8], будет $\overrightarrow{X'Y'} = \overrightarrow{XY''}$. Итак, $\overrightarrow{X'Y'}$ — это составляющая вектора $\overrightarrow{XY}$ на плоскости α при его разложении вдоль прямой a и вдоль плоскости α. Вместе с тем проведенное построение этой составляющей представляет собой не что иное, как параллельное проектирование точек X, Y на плоскость α вдоль прямой a. Теорема 2 доказана. □

Из этой теоремы следует

Теорема 3. *При параллельном проектировании сохраняются линейные соотношения между векторами, т. е. если $\overrightarrow{AB} = p\overrightarrow{CD} + q\overrightarrow{EF}$, то $\overrightarrow{A'B'} = p\overrightarrow{C'D'} + q\overrightarrow{E'F'}$. Тем самым равным векторам отвечают равные (например $p = 1$, $q = 0$), сумме — сумма, произведению на число — произведение на то же число.*

Доказательство. Основное свойство составляющих любых векторов состоит в том, что линейным соотношениям между векторами соответствуют такие же соотношения между их составляющими. Это заключает в себе сказанное в теореме 3. □

Теперь свойства, указанные в теореме 1, вытекают из теоремы 3 благодаря следующему предложению.

Лемма. *Если при некотором отображении сохраняются линей-*

---

[8]Прямые XX', YY' параллельны и, стало быть, лежат в одной плоскости γ. Эта плоскость пересекает α и β по параллельным прямым. Итак $XX' || Y''Y'$, $XY'' || X'Y'$. Поэтому отрезки XY'' и $X'Y'$ — стороны параллелограмма, и, значит, $\overrightarrow{XY''} = \overrightarrow{X'Y'}$.

ные соотношения векторов, то при этом отображении выполняются свойства, указанные в теореме 1, *для отрезков и прямых, не отображающихся в точку*: 1) *отрезки* (*прямые*) *отображаются на отрезки* (*прямые*), 2) *параллельные — на параллельные*, 3) *отношения параллельных отрезков сохраняются.*

Доказательство. Отрезок AB заполняется концами векторов $\overrightarrow{AX} = p\overrightarrow{AB}$ с $p \in [0,\ 1]$. Параллельные отрезки соответствуют параллельным (коллинеарным) векторам. Отношение таких отрезков — это отношение этих векторов, точнее — если $\overrightarrow{AB} = p\overrightarrow{CD}$, то $AB = |p|CD$. Поэтому вместе с сохранением линейных соотношений между векторами отрезки отображаются на отрезки, параллельные — на параллельные, и отношения отрезков сохраняются, что и требовалось доказать. □

Вместе с этим доказана и теорема 1. □

Проектирование с плоскости на плоскость.

Теорема 4. *При параллельном проектировании с плоскости на плоскость линейные соотношения между векторами сохраняются и новые не появляются* (подразумевается, конечно, что прямая, вдоль которой происходит проектирование, пересекает обе плоскости).

Доказательство. Проектирование с одной плоскости α на другую β в направлении прямой a, пересекающей обе плоскости, устанавливает взаимно однозначное соответствие между точками этих плоскостей, так что как плоскость α проектируется на β вдоль прямой a, так и β проектируется на α вдоль той же прямой. Поэтому, согласно теореме 3, линейным соотношениям между векторами на одной плоскости соответствуют такие же соотношения на другой, и обратно.

Стало быть, если бы при проектировании с плоскости α на β появлялись новые линейные соотношения между векторами, то при обратном проектировании — с β на α — они должны были бы дать линейные соотношения на плоскости α, но, значит, они не были «новыми». Следовательно, никаких новых линейных соотношений между векторами при проектировании с одной плоскости на другую не появляется, что и требовалось доказать. □

Замечание. Пусть $\boldsymbol{u}_\alpha$ — вектор на плоскости α (параллельный α). Проектируя вдоль прямой a на плоскость β, мы разлагаем его на составляющие: $\boldsymbol{u}_\beta$ — на β и $\boldsymbol{u}_a$ — вдоль a, так что $\boldsymbol{u}_\alpha = \boldsymbol{u}_\beta + \boldsymbol{u}_a$. Отсюда $\boldsymbol{u}_\beta = \boldsymbol{u}_\alpha - \boldsymbol{u}_a$, т. е. мы так же разлагаем вектор $\boldsymbol{u}_\beta$ на плоскости β (рис. 58) и так же, проектируя вдоль a с β на α, разлагаем любой

вектор $\boldsymbol{v}_\beta$ плоскости β: $\boldsymbol{v}_\beta = \boldsymbol{v}_\alpha + \boldsymbol{v}_a$. Ясно, что эти разложения равносильны.

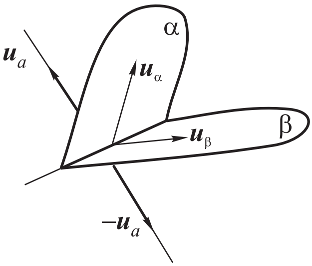


Рис. 58

Спроектировав плоскость α на β вдоль прямой a, можно затем спроектировать β на α вдоль какой-либо прямой b, или спроектировать β на γ, а потом — γ на α, и т. п. Так комбинируя проектирования с одной плоскости на другую, можно получать отображения данной плоскости α на себя. При этом, поскольку при каждом проектировании с плоскости на плоскость не исчезают и не появляются линейные зависимости векторов, то и при получающемся отображении плоскости на себя линейные соотношения векторов сохраняются и новые не появляются.

Отображения, обладающие таким свойством, называются *аффинными*. Их мы дальше и рассмотрим.

§ 2. Аффинные отображения и аффинная геометрия

Определения. Отображение какой бы то ни было пространственной или плоской фигуры в пространство называется *аффинным*, если оно сохраняет линейные соотношения между векторами и не вводит новых.

Первое условие означает, что если в фигуре есть такие точки A, B, $\ldots$, G, H, что при некоторых числах $p, \ldots, s$

$$p\overrightarrow{AB} + \ldots + s\overrightarrow{GH} = \mathbf{0},$$

то для образов этих точек A', B', $\ldots$, G', H'

$$p\overrightarrow{A'B'} + \ldots + s\overrightarrow{G'H'} = \mathbf{0}.$$

Второе условие означает, что и обратно: если в фигуре F', полученной из F аффинным отображением, есть такие точки P', Q', $\ldots$, X', Y', что при некоторых числах $k, \ldots, n$

$$k\overrightarrow{P'Q'} + \ldots + n\overrightarrow{X'Y'} = \mathbf{0},$$

то и в исходной фигуре F для соответствующих точек P, Q, $\ldots$, X, Y (прообразов P', Q' $\ldots$ и т. д.) выполняется аналогичное равенство.

Свойство фигуры называется *аффинным*, если оно сохраняется при любых аффинных отображениях.

Аффинной геометрией называется та часть геометрии, в которой изучаются аффинные свойства фигур.

Из определения аффинного отображения непосредственно следует, что всякое свойство фигур, которое может быть выражено через линейные соотношения между векторами, является аффинным.

Из теоремы 4 предыдущего параграфа следует, что параллельное проектирование с плоскости на плоскость является аффинным отображением. Как отмечено в конце предыдущего параграфа, всякое отображение, получающееся в результате ряда проектирований с плоскости на плоскость, является аффинным. Оно относится к плоским фигурам, но, конечно, возможны разнообразные аффинные отображения пространственных фигур.

Простым примером аффинного свойства фигуры может служить наличие у нее центра симметрии. Центр симметрии O характеризуется тем, что для всякой точки M фигуры есть такая M', что $\overrightarrow{OM'} = -\overrightarrow{OM}$. Это свойство выражено через линейное соотношение векторов и, стало быть, аффинно.

Общие свойства аффинных отображений.

Теорема 1. *При аффинном отображении отрезки отображаются на отрезки, причем параллельные — на параллельные с сохранением отношений; плоские фигуры отображаются на плоские фигуры; прямые — на прямые, плоскости — на плоскости, полуплоскости — на полуплоскости.*

Доказательство. То, что при аффинном отображении отрезки отображаются на отрезки, параллельные — на параллельные с сохранением отношения, следует из определения аффинного отображения в силу леммы, доказанной в предыдущем параграфе.

Плоскость есть множество концов векторов, отложенных от одной точки и выражающихся через два неколлинеарных вектора. Поэтому плоскость отображается на плоскость и, аналогично, плоские фигуры — на плоские фигуры. (Ср. с теоремой 2 ниже.) □

То, что отношения параллельных отрезков сохраняются, означает, что у всех параллельных друг другу отрезков длины изменяются в одно и то же число раз; другими словами, все эти отрезки претерпевают одинаковое растяжение или сжатие. Таким образом, *при аффинном преобразовании в каждом направлении происходит равномерное растяжение или сжатие.*

Из доказанной теоремы следует, что все свойства фигур, основанные на отношениях параллельных отрезков (в частности, отрезков, лежащих на одной прямой), являются аффинными. Теоремы, их выражающие, принадлежат аффинной геометрии. Примером может служить теорема о медианах треугольника, по которой они пересекаются в одной точке и делят друг друга в отношении $1:2$.

Теорема 2. *Аффинное отображение всякой фигуры можно распространить на все пространство, причем единственным образом, если в фигуре есть три некомпланарных вектора. Если же все векторы в ней компланарны, но есть неколлинеарные, то аффинное отображение однозначно распространяется на определяемую ими плоскость.*

Доказательство. Пусть в фигуре G есть такие точки A, ..., F, что векторы $\overrightarrow{AB}$, $\overrightarrow{CD}$, $\overrightarrow{EF}$ некомпланарны. Тогда всякую точку M в пространстве можно задать вектором

$$\overrightarrow{AM} = x\overrightarrow{AB} + y\overrightarrow{CD} + z\overrightarrow{EF};$$

x, y, z — координаты точки M относительно трехвекторника с началом A и векторами, равными $\overrightarrow{AB}$, $\overrightarrow{CD}$, $\overrightarrow{EF}$.

При аффинном отображении фигуры G эти векторы отобразятся в независимые векторы, и будет

$$\overrightarrow{A'M'} = x\overrightarrow{A'B'} + y\overrightarrow{C'D'} + z\overrightarrow{E'F'}.$$

Этим однозначно определено отображение всего пространства на себя. В частности, векторы $\overrightarrow{XY}$ в самой фигуре G выражаются через $\overrightarrow{AB}$, $\overrightarrow{CD}$, $\overrightarrow{EF}$. То есть фигура G преобразуется вместе с пространством.

Если в фигуре нет трех независимых векторов, т. е. она содержится в некоторой плоскости α, то при аффинном отображении это свойство сохраняется; образ фигуры будет содержаться в некоторой плоскости β. Чтобы распространить отображение на все пространство, можно добавить к фигуре точку, не лежащую в плоскости α, и задать ее образ, лишь бы он не лежал в плоскости β. $\square$

Теорема 3. *Композиция аффинных отображений есть аффинное отображение. Всякое аффинное отображение обратимо, и отображение, обратное аффинному, аффинно.*

Доказательство. Первое утверждение очевидно, так как если отображения что-то сохраняют, то и композиция их это сохраняет. При аффинном отображении никакие две точки не отображаются на одну,

потому что если две точки A, B отображаются на одну A', то, значит, вектор $\overrightarrow{AB}$ отображается на нулевой, а это исключено определением аффинного отображения. Следовательно, аффинное отображение взаимно однозначно и тем самым обратимо.

Если бы при обратном отображении исчезала какая-нибудь линейная связь между векторами, то это значило бы, что при прямом отображении она появляется. А это противоречит второму условию в определении аффинного отображения. А если бы при обратном отображении появлялась какая-нибудь линейная связь векторов, то это означало бы, что при прямом отображении она исчезает вопреки определению аффинного отображения. Теорема 3 доказана. □

Аффинное отображение плоскости на себя, как и пространства на себя, называется *аффинным преобразованием* плоскости, и соответственно — пространства.

Из теоремы 3 непосредственно следует

Теорема 4. *Аффинные преобразования плоскости, как и аффинные преобразования пространства, образуют группу.* □

Об определении аффинных отображений. Определение аффинного отображения можно несколько упростить. Первое его требование — сохранение линейных связей векторов — остается, но второе — то, что не появляется новых таких связей, — можно заменить более простым: сохраняется независимость векторов, т. е. образы неколлинеарных векторов неколлинеарны, некомпланарных — некомпланарны. То, что это условие выполнено, если вообще не появляется новых линейных связей, очевидно. То, что верно также обратное, доказывает следующая лемма.

Лемма. *Если при отображении, сохраняющем линейные соотношения между векторами, сохраняются неколлинеарность и некомпланарность векторов, то вообще не появляется новых линейных соотношений.*

Доказательство. Пусть, например, в отображаемой фигуре F есть векторы $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$, причем $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ независимы и

$$p\boldsymbol{a} + q\boldsymbol{b} + r\boldsymbol{c} + s\boldsymbol{d} = \mathbf{0}.$$

При отображении по первому условию получим векторы с такой же зависимостью

$$p\boldsymbol{a}' + q\boldsymbol{b}' + r\boldsymbol{c}' + s\boldsymbol{d}' = \mathbf{0}. \tag{1}$$

Допустим, появилась также другая зависимость

$$p_1\boldsymbol{a}' + q_1\boldsymbol{b}' + r_1\boldsymbol{c}' + s_1\boldsymbol{d} = \mathbf{0}. \tag{2}$$

Умножая это равенство на s, а предыдущее — на s_1 и вычитая, исключаем вектор $\boldsymbol{d}'$ и получаем формулу вида

$$p_2\boldsymbol{a}' + q_2\boldsymbol{b}' + r_2\boldsymbol{c}' = \boldsymbol{0},$$

где $p_2 = ps_1 - p_1 s$ и т. д.

Если векторы $\boldsymbol{a}'$, $\boldsymbol{b}'$, $\boldsymbol{c}'$ независимы, как и $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, то должно быть

$$p_2 = q_2 = r_2 = 0,$$

т. е.

$$p_1 : p = s_1 : s, \quad q_1 : q = s_2 : s, \quad r_1 : r = s_1 : s.$$

А это значит, что равенство (2) равносильно (1), т. е. вовсе не представляет новую зависимость векторов.

Стало быть, если некомпланарные векторы остаются некомпланарными, то новых зависимостей не появляется.

В случае плоскости все векторы компланарны. Тогда аналогично, — рассматривая зависимость вида $p\boldsymbol{a} + q\boldsymbol{b} + r\boldsymbol{c} = \boldsymbol{0}$, — доказываем, что если неколлинеарные векторы остаются неколлинеарными, то новых зависимостей не появляется. □

Хотя мы и упростили несколько определение аффинного отображения, оно, тем не менее, остается довольно сложным, так как основано на понятиях векторной алгебры и недостаточно геометрично. Поэтому существенно определить аффинные отображения более геометричными и менее сложными условиями. Это можно сделать, если рассматривать аффинные отображения не любых фигур.

Сформулируем два рода таких условий; первые относятся к такому случаю, когда отображению подвергается область, во втором речь идет об отображении всей плоскости или всего пространства.

Теорема 5. *Отображение плоской или пространственной области, при котором параллельные отрезки отображаются на параллельные отрезки, является аффинным.*

Теорема 6. *Отображение плоскости или пространства на себя аффинно, если при нем всякие три точки, лежащие на одной прямой, переходят в точки, опять лежащие на одной прямой.*

Или: взаимно однозначное отображение плоскости или пространства, при котором прямые отображаются на прямые, аффинно.

При таком отображении пересекающиеся прямые переходят в пересекающиеся, и поэтому, как можно доказать, плоскость отображается на плоскость. Поэтому параллельные прямые отображаются на параллельные (в силу взаимной однозначности). Но чтобы свести вывод к

теореме 5, нужно доказать, что отрезки отображаются на отрезки, а это не просто.

Теоремы 5 и 6 мы доказывать не будем. □

§ 3. Разложение аффинных отображений на простейшие

Наложение как аффинное отображение.

Теорема 1. *Всякое наложение есть аффинное отображение.*

Доказательство. Наложение любой фигуры можно распространить на все пространство; поэтому можно рассматривать наложение, определенное во всем пространстве.

Согласно определению аффинных отображений мы должны убедиться в том, что наложение сохраняет линейные соотношения между векторами и не вводит новых. При этом достаточно ограничиться проверкой того, что линейные соотношения сохраняются. Тогда в том, что новых соотношений не появится, убеждаемся, применяя доказанное к обратному наложению.

Доказательство разбивается на три части. Сначала мы проверим, что образы равных направленных отрезков равны. Затем — что произведение вектора на число переходит в такое же произведение. И, наконец, что образ суммы векторов равен сумме их образов.

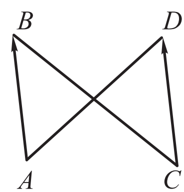


Рис. 59

Векторное равенство $\overrightarrow{AB} = \overrightarrow{CD}$ равносильно тому, что у отрезков AD, BC середина общая (рис. 59). При наложении середина переходит в середину, поэтому и равенство векторов сохраняется. Это позволяет заменять вектор любым равным — образ его также заменяется равным.

В равенстве $p\overrightarrow{AB} = \overrightarrow{CD}$ можно вектор $\overrightarrow{CD}$ заменить равным, отложенным от точки A, так что $p\overrightarrow{AB} = \overrightarrow{AE}$. Это равенство означает, во-первых, что $|p||AB| = |AE|$; а соотношение длин при наложении сохранится. Во-вторых, если $p < 0$, то точка A лежит на отрезке BE; если же $0 < p < 1$, то точка E лежит на отрезке AB, и если $p > 1$, то B на AE. Все эти соотношения при наложении сохраняются, так как отрезки отображаются на отрезки и длины их сохраняются.

В сумме $\overrightarrow{AB} + \overrightarrow{CD} = \overrightarrow{EF}$ можно заменять $\overrightarrow{CD}$ и $\overrightarrow{EF}$ равными векторами, отложенными от точек B и A, так что будет $\overrightarrow{AB} + \overrightarrow{BG} = \overrightarrow{AG}$. А такое равенство очевидным образом сохраняется вообще при любом

отображении, так как оно не выражает ничего большего, как то, что имеются точки A, B, G. $\square$

Представление аффинного отображения через растяжения и наложения. Простейшим аффинным преобразованием плоской фигуры, не сводящимся к наложению, является равномерное растяжение или сжатие к прямой по перпендикулярному ей направлению. Т. е. это такое преобразование, при котором все расстояния от данной прямой изменяются в одно и то же число раз и каждая точка остается при этом на том же перпендикуляре к этой прямой с той же от нее стороны. Если эту прямую принять за ось x прямоугольной системы координат, то преобразование представится так, что каждая точка $(x,\ y)$ переходит

а *б* *в*

Рис. 60

в такую точку (x', y'), что

$$x' = x, \quad y' = ky, \tag{1}$$

где $k > 0$ одно и то же для всех точек. Это k — «коэффициент растяжения», причем если $k > 1$, то происходит в самом деле растяжение в k раз, если же $k < 1$, то сжатие в $1/k$ раз (случай $k = 1$, как это принято, формально не исключается). См. рис. 60, *а, б, в.*

Нетрудно доказать, что такое преобразование в самом деле аффинно, т. е. сохраняет линейные связи векторов и не вводит новых. Действительно, преобразование (1) очевидно сохраняет по отдельности линейные связи координат x и координат y (x вовсе не меняется, а у всех y общий множитель, который сократится во всякой формуле

$y_3 = \lambda_1 y_1 + \lambda_2 y_2$). Линейные же связи векторов равносильны таким же линейным связям их координат. Таким образом, эти связи сохраняются. Обратное преобразование имеет такой же вид: $x = x'$; $\quad y = \frac{1}{k} y'$. Поэтому оно также сохраняет линейные связи векторов. А это значит, что прямое преобразование не вводит новых линейных связей. Итак, преобразование (1) аффинно.

Роль таких простейших преобразований показывает следующая теорема.

Теорема 2. *Всякое аффинное преобразование на плоскости представляет собой композицию наложения и двух растяжений во взаимно перпендикулярных направлениях от двух взаимно перпендикулярных прямых.*

Поскольку наложение не изменяет ни формы, ни размеров фигуры, то теорема означает, что изменение их, вызываемое аффинным преобразованием (рис. 61), сводится к двум сжатиям или растяжениям (причем, конечно, одно из них может отсутствовать).

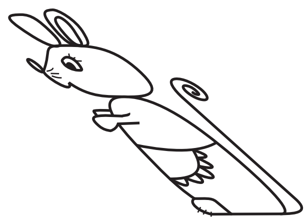


Рис. 61

Если оси x, y направлены по направлениям растяжения (сжатия), то преобразование, слагающееся из этих двух растяжений, состоит в том, что всякая точка $(x,\ y)$ переходит в точку (x', y') с

$$x' = k_1 x, \quad y' = k_2 x. \tag{2}$$

Теорему можно так и формулировать: *при всяком аффинном преобразовании можно выбрать прямоугольные координаты x, y так, что преобразование представится формулами* (2) *с последующим наложением* (впрочем, можно сначала произвести необходимое наложение, а потом растяжения (сжатия)).

В пространстве простейшим аффинным преобразованием, не сводящимся к наложению, является растяжение или сжатие к плоскости по перпендикулярному ей направлению. Все расстояния от данной плоскости изменяются в одно и то же число раз, и каждая точка остается при этом на том же перпендикуляре к этой плоскости с той же от нее стороны. Если ввести прямоугольные координаты так, чтобы эта плоскость была плоскостью xy ($z = 0$), то указанное преобразование представится так, что каждая точка $(x,\ y,\ z)$ переходит в такую точку $(x',\ y',\ z')$, что

$$x' = x,\ y' = y,\ z' = kz. \tag{3}$$

То, что такое преобразование в самом деле аффинно, доказывается буквально так же, как в случае плоскости, и аналогично теореме 2 выполняется.

Теорема 3. *Всякое аффинное преобразование в пространстве представляет собой композицию наложения и трех растяжений (сжатий) во взаимно перпендикулярных направлениях от трех взаимно перпендикулярных плоскостей.*

Если оси x, y, z направлены по направлениям указанных в теореме растяжений, то преобразование, состоящее из этих растяжений, представляется как такое, при котором каждая точка $(x,\ y,\ z)$ переходит в такую точку $(x',\ y',\ z')$, что

$$x' = k_1 x, \quad y' = k_2 y, \quad z = k_3 z. \tag{4}$$

Теорему можно так и формулировать: *при всяком аффинном преобразовании в пространстве можно ввести прямоугольные координаты так, что преобразование представится формулами* (4) *«с точностью до наложения», т. е. с добавлением наложения.*

Доказательства теорем 2, 3. Как было указано в предыдущем параграфе, при аффинном отображении в каждом направлении происходит равномерное растяжение. Длины всех параллельных друг другу отрезков изменяются в одно и то же число раз, умножаются на одно и то же число — коэффициент растяжения (возможно, < 1).

Теорему 2 можно дополнить:

Теорема 2а. *Растяжения в двух взаимно перпендикулярных направлениях, которые согласно теореме* 2, *дают вместе с наложением данное аффинное преобразование, происходят в тех направлениях, в которых коэффициент растяжения имеет наименьшее и наибольшее значения.*

С указанным экстремальным свойством этих растяжений и связано доказательство теоремы 2. Оно основано на свойстве наименьшего растяжения, выраженном в следующей лемме. В ней имеется в виду аффинное преобразование плоскости.

Лемма 1. *Прямые, перпендикулярные направлению наименьшего растяжения, остаются ему перпендикулярными.*

Доказательство. Пусть имеется некоторое аффинное преобразование. В направлении каждой прямой происходит равномерное растяжение. Пусть a — проходящая через точку O прямая, растяжение вдоль которой наименьшее (рис. 62). Возьмем на ней какую-нибудь точку A, отличную от O, и проведем через нее прямую b, перпендикулярную a. В результате произведенного аффинного преобразования

точка O перейдет в O', прямая a перейдет в какую-то прямую a', точка A — в точку A', прямая b — в прямую b'. Докажем, что $b' \perp a'$.

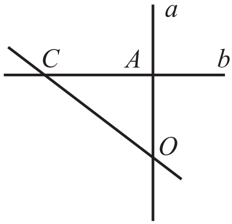


Рис. 62

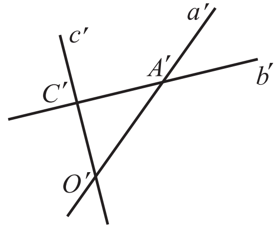


Рис. 63

Допустим противное, и пусть c' — прямая, проходящая через O' перпендикулярно b', а c — та прямая, которая перешла в c' при рассматриваемом преобразовании. Пусть C — точка пересечения прямой c с прямой b, и C' — образ этой точки — точка пересечения прямой c' с прямой b'. (рис. 63).

Таким образом, $O'C' \perp b'$, и, стало быть,

$$O'C' < O'A'.$$

Вместе с тем $OA \perp b$, и потому

$$OC > OA.$$

Из этих двух неравенств следует, что

$$\frac{O'C'}{OC} < \frac{O'A'}{OA}.$$

Это неравенство означает, что растяжение в направлении прямой c меньше, чем в направлении прямой a. Но это противоречит выбору прямой a как такой, в направлении которой растяжение наименьшее. Следовательно, предположение, что прямая b', в которую перешла b, не перпендикулярна прямой a', неверно, и, значит, $b' \perp a'$, что и требовалось доказать. $\square$

Теперь укажем кратко доказательство теоремы 2.

Пусть дано некоторое аффинное преобразование плоскости. Добавив к нему подходящий перенос, обеспечим, что некоторая точка будет неподвижной. Затем, добавив подходящий поворот, обеспечим,

что прямая a, проходящая через O в направлении наименьшего растяжения, остается на месте, как и ее лучи с началом O (прямая не переворачивается). Тогда согласно доказанной лемме прямая b, перпендикулярная a, тоже останется на месте, и если ее лучи с началом O меняются местами, то, добавив отражение в прямой a, обеспечим, что они остаются на месте.

Итак, добавив к данному аффинному преобразованию подходящее наложение, мы получаем аффинное преобразование, при котором прямые a и b и их лучи переходят в себя. При аффинном преобразовании параллельные прямые переходят в параллельные, поэтому прямые, параллельные прямым a и b, остаются им параллельными и только перемещаются согласно растяжениям вдоль этих прямых. А это и значит, что аффинное преобразование состоит из растяжений вдоль прямых a и b.

Тем самым исходное преобразование является композицией этих растяжений с наложением (обратным тому, каким мы привели к неподвижности прямые a и b). Теорема 2 доказана. □

Из доказательства следует, что одно из растяжений, из которых слагается здесь аффинное преобразование,— это растяжение с наименьшим коэффициентом, как и сказано в теореме 2а. В том, что другое растяжение оказывается наибольшим, можно убедиться, выразив растяжение в любом данном направлении через растяжение по двум взаимно перпендикулярным прямым a и b. Но можно сослаться на обратное отображение. При обратном отображении прямые a и b будут также оставаться на месте, но наименьшее растяжение соответствует наибольшему в прямом отображении. Следовательно, это наибольшее растяжение происходит как раз вдоль прямой b. □

Теорема 3 об аффинном преобразовании пространства доказывается аналогично. Начинаем с того, что подобно доказанной лемме 1 доказываем:

Лемма 2. *При аффинном преобразовании пространства все плоскости, перпендикулярные направлению наименьшего растяжения, остаются ему перпендикулярными.*

Доказательство совершенно аналогично доказательству леммы 1. □

Далее замечаем, что в параллельных плоскостях произведены одинаковые аффинные преобразования. В плоскости, перпендикулярной прямой a, вдоль которой происходит наименьшее растяжение, аффинное преобразование представляется согласно теореме 2.

Так придем к доказательству теоремы 3. Читатель сам проведет это доказательство. □

Замечание. В доказательствах теорем 2, 3 допущен пропуск: нужно было доказать, что при аффинном преобразовании всегда есть направление наименьшего растяжения. Это доказывается обычным приемом анализа, если доказать, что коэффициент растяжения в каждом направлении является непрерывной функцией направления. На доказательстве мы не останавливаемся.

§ 4. Представление аффинных отображений и наложений в координатах

Теорема 1. *Аффинное преобразование представляется в аффинных координатах линейными формулами с определителем, отличным от нуля. То есть в случае плоскости координаты точки $M'(x',\ y')$, на которую отображается точка $M(x,\ y)$, выражаются формулами вида*

$$\begin{aligned} x' &= a_{11}x + a_{12}y + b_1, \\ y' &= a_{21}x + a_{22}y + b_2, \end{aligned} \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0. \tag{1}$$

В пространстве же, если точка $M(x,\ y,\ z)$ отображается на $M'(x',\ y',\ z')$, то

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + b_3, \end{aligned} \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0. \tag{2}$$

Доказательство. Для плоскости. Пусть на плоскости введена система координат с началом O и основными векторами $\boldsymbol{e}_1,\ \boldsymbol{e}_2$. Пусть произведено аффинное преобразование, при котором начало O отобразилось на точку $O'(b_1,\ b_2)$, а векторы $\boldsymbol{e}_1,\ \boldsymbol{e}_2$ — на $\boldsymbol{e}_1',\ \boldsymbol{e}_2'$. Возьмем произвольную точку $M(x_1,\ x_2)$, и пусть ее радиус-вектор будет

$$\overrightarrow{OM} = \boldsymbol{e}_1 x_1 + \boldsymbol{e}_2 x_2. \tag{3}$$

При аффинном отображении линейные соотношения сохраняются. Поэтому если $M'(x_1',\ x_2')$ — та точка, на которую отобразилась точка M, то

$$\overrightarrow{O'M'} = \boldsymbol{e}_1' x_1 + \boldsymbol{e}_2' x_2. \tag{4}$$

Вместе с тем радиус-вектор этой точки M' в основных векторах представляется формулой

$$\overrightarrow{OM'} = \boldsymbol{e}_1 x_1' + \boldsymbol{e}_2 x_2'. \tag{5}$$

И так как

$$\overrightarrow{OM'} = \overrightarrow{OO'} + \overrightarrow{O'M'}; \quad \overrightarrow{OO'} = \boldsymbol{b} = \boldsymbol{e}_1 b_1 + \boldsymbol{e}_2 b_2,$$

то из (4) и (5) получаем

$$\overrightarrow{OM'} = \boldsymbol{e}_1 x_1' + \boldsymbol{e}_2 x_2' = \boldsymbol{e}_1' x_1 + \boldsymbol{e}_2' x_2 + \boldsymbol{e}_1 b_1 + \boldsymbol{e}_2 b_2. \tag{6}$$

Векторы $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$ выражаются через векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$:

$$\boldsymbol{e}_1' = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2, \quad \boldsymbol{e}_2' = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2.$$

Подставив эти выражения в (3), получим

$$\boldsymbol{e}_1 x_1' + \boldsymbol{e}_2 x_2' = (a_{11}x_1 + a_{12}x_2 + b_1)\boldsymbol{e}_1 + \ldots.$$

Так как разложение вектора по двум независимым векторам $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ единственное, то множители при $\boldsymbol{e}_1$ и $\boldsymbol{e}_2$ в обеих частях равны, поэтому получаем формулы

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + b_1, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + b_2. \end{aligned} \tag{7}$$

Аффинное отображение обратимо, поэтому как x_1', x_2' выражаются через x_1, x_2, так и обратно: x_1, x_2 должны выражаться через x_1', x_2'; т. е. уравнения (1) должны быть однозначно разрешимы относительно x_1, x_2. Необходимым условием такой разрешимости системы уравнений является неравенство нулю ее определителя; так что определитель в (7) отличен от нуля, что и требовалось доказать.

Доказательство для пространства отличается от приведенного доказательства для плоскости только тем, что появляется третья координата вместе с третьим базисным вектором. Читатель сам изложит для себя получающийся таким образом вывод. □

Представление наложения в прямоугольных координатах. Наложение представляет собою аффинное преобразование, поэтому оно представляется в координатах таким же образом. Однако если пользоваться прямоугольными координатами, то коэффициенты в формулах (1), (2) обладают особыми свойствами, такими же, как в преобразовании прямоугольных координат. Именно в матрице (таблице) этих коэффициентов сумма их квадратов в каждой строке, как и в каждом столбце, равна единице, а суммы произведений одноименных

коэффициентов из двух разных строк, как и двух разных столбцов, равна нулю. Для двух переменных — формулы (1) — сказанное выражается так: для строк

$$a_{11}^2 + a_{12}^2 = a_{21}^2 + a_{22}^2 = 1, \quad a_{11}a_{21} + a_{12}a_{22} = 0$$

и аналогично для столбцов. Для трех переменных — формулы (2) — для первых строк

$$a^2 11 + a_{12}^2 + a_{13}^2 = 1, \quad a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} = 0$$

и аналогично для других строк и для столбцов.

Матрица с такими свойствами называется *ортогональной,* а преобразование переменных с такой матрицей называется *ортогональным.*

Это свойство матрицы коэффициентов обусловлено их геометрическим смыслом, который мы сейчас выясним, доказав следующую теорему.

Теорема 2. *Наложение представляется в прямоугольных координатах ортогональным преобразованием. Его коэффициенты — это косинусы углов между исходными базисными векторами и теми, в какие они преобразуются при наложении.*

Доказательство. Проведем его для случая пространства. В пространстве формула (6) для радиус-вектора $\overrightarrow{OM'}$ примет вид

$$\boldsymbol{e}_1 x_1' + \boldsymbol{e}_2 x_2' + \boldsymbol{e}_3 x_3' = \boldsymbol{e}_1' x_1 + \boldsymbol{e}_2' x_2 + \boldsymbol{e}_3' x_3 + \boldsymbol{e}_1 b_1 + \boldsymbol{e}_2 b_2 + \boldsymbol{e}_3 b_3. \qquad (8)$$

Векторы $\boldsymbol{e}_i$, как и $\boldsymbol{e}_i'$, единичные и взаимно ортогональные:

$$\boldsymbol{e}_1^2 = \boldsymbol{e}_2^2 = \boldsymbol{e}_3^2 = 1, \quad \boldsymbol{e}_1\boldsymbol{e}_2 = \boldsymbol{e}_2\boldsymbol{e}_3 = \boldsymbol{e}_3\boldsymbol{e}_1 = 0.$$

Поэтому, умножая (8) на $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ и $\boldsymbol{e}_3$, получим

$$\begin{aligned}
x_1' &= (\boldsymbol{e}_1\boldsymbol{e}_1')x_1 + (\boldsymbol{e}_1\boldsymbol{e}_2')x_2 + (\boldsymbol{e}_1\boldsymbol{e}_3')x_3 + b_1,\\
x_2' &= (\boldsymbol{e}_2\boldsymbol{e}_1')x_1 + (\boldsymbol{e}_2\boldsymbol{e}_2')x_2 + (\boldsymbol{e}_2\boldsymbol{e}_3')x_3 + b_2,\\
x_3' &= (\boldsymbol{e}_3\boldsymbol{e}_1')x_1 + (\boldsymbol{e}_3\boldsymbol{e}_2')x_2 + (\boldsymbol{e}_3\boldsymbol{e}_3')x_3 + b_3.
\end{aligned}$$

Стоящие здесь скалярные произведения и есть коэффициенты a_{ij} в формуле. Так как векторы $\boldsymbol{e}_i'$ единичные и взаимно ортогональные, то скалярные произведения $\boldsymbol{e}_1\boldsymbol{e}_i'$, $\boldsymbol{e}_2\boldsymbol{e}_i'$, $\boldsymbol{e}_3\boldsymbol{e}_i'$ суть не что иное, как координаты векторов $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ относительно векторов $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$, $\boldsymbol{e}_3'$. А так

как сами $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ единичные и взаимно ортогональные, то для них сумма квадратов координат равна единице, а сумма произведений их одноименных координат, т. е. их скалярное произведение, равна нулю.

Аналогично видно, что коэффициенты при x_1 — это координаты вектора $\boldsymbol{e}_1$ относительно векторов $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$, $\boldsymbol{e}_3'$, и тот же смысл имеют коэффициенты при x_2 и x_3. А так как векторы $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$, $\boldsymbol{e}_3'$ единичные и взаимно ортогональные, то и для коэффициентов по столбцам получается то же: суммы квадратов равны единице, а суммы попарных произведений из разных столбцов равны нулю.

Итак, доказано, что матрица коэффициентов a_{ik} ортогональная.

Так как векторы $\boldsymbol{e}_i$, $\boldsymbol{e}_k'$ единичные, то скалярное произведение их — это косинус угла между ними:

$$a_{ik} = \boldsymbol{e}_i \boldsymbol{e}_k' = \cos\angle(\boldsymbol{e}_i,\ \boldsymbol{e}_k'),$$

что и требовалось доказать. □

Замечание о гладких отображениях. Аффинные отображения играют важную роль не только сами по себе, но и потому, что всякое «не слишком скверное», «гладкое» отображение является аффинным «в бесконечно малом», т. е. во всякой малой области с точностью до величин, пренебрежимо малых в сравнении с размерами самой области. Так, всякая упругая деформация представляет для малого куска деформируемого тела его аффинное отображение с точностью тем большей, чем меньше кусок. Поэтому для такого куска деформация представляет собой, в согласии с теоремой 3, § 3, сочетание перемещения с растяжением по трем взаимно перпендикулярным направлениям, тем точнее, чем меньше рассматриваемый кусок. При упругой деформации скажем, плоской пластинки или пленки из резины каждый малый ее кусок испытывает в соответствии с теоремой 2, § 3 некоторое перемещение и растяжение по двум взаимно перпендикулярным направлениям (опять-таки, тем точнее, чем меньше кусок).

Перемещение не приводит к деформации, так что деформация в собственном смысле слова всегда сводится в малых частях деформируемого тела к растяжениям или сжатиям по трем или двум взаимно перпендикулярным направлениям. Одно из них — направление наименьшего растяжения, другое — направление наибольшего растяжения.

Математически это представляется для плоских фигур так.

Пусть G — область на плоскости α, и пусть задано ее отображение f в плоскость β (может быть, совпадающую с α). Пусть в плоскостях α, β введены координаты x, y и u, v — прямоугольные или, вообще, аффинные. Отображение f представляется так, что каждой точке $M(x,\ y)$

из области G ставится в соответствие точка $N(u,\ v)$:

$$u = u(x,\ y), \quad v = v(x,\ y).$$

Обратное отображение представляется аналогично:

$$x = x(u, v), \quad y = y(u,\ v).$$

Пусть фигурирующие здесь функции дифференцируемы. Тогда обозначая через $u_x,\ \ldots,\ v_y$ частные производные от u, v по x и y, получаем дифференциалы:

$$du = u_x dx + u_y dy, \quad dv = v_x dx + v_y dy, \tag{9}$$

и аналогично dx, dy выражаются через du, dv. Поэтому формулы (9) обратимы, так что определитель не равен нулю:

$$\begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} \neq 0. \tag{10}$$

Возьмем точку $(x_0,\ y_0)$ из области G и положим

$$u_0 = u(x_0,\ y_0), \quad v_0 = v(x_0,\ y_0),$$
$$u - u_0 = \Delta u,\ v - v_0 = \Delta v, \quad x - x_0 = \Delta x,\ y - y_0 = \Delta y.$$

Согласно известным выводам дифференциального исчисления приращение функций u, v будет

$$\begin{aligned} \Delta u &= u_x \Delta x + u_y \Delta y + \varepsilon_1 \sqrt{\Delta x^2 + \Delta y^2}, \\ \Delta v &= v_x \Delta x + v_y \Delta y + \varepsilon_2 \sqrt{\Delta x^2 + \Delta y^2}, \end{aligned} \tag{11}$$

где $u_x,\ \ldots,\ v_y$ — частные производные в точке $(x_0,\ y_0)$, т. е. некоторые числа, а ε_1, $\varepsilon_2 \to 0$ при

$$\sqrt{\Delta x^2 + \Delta y^2} \to 0.$$

Если пренебречь членами с ε_1, ε_2, то формулы (11) ввиду условия (10) представляют аффинное отображение.

В трехмерном случае вывод будет аналогичным.

Глава IV

ПРОЕКТИВНАЯ ГЕОМЕТРИЯ

§ 1. Проективная плоскость и проективная геометрия

Проективная геометрия возникла из изучения проектирования фигур на плоскость из какой-либо точки — центра проекции. При проектировании с плоскости α на плоскость α' из центра O точке $M \in \alpha$ сопоставляется точка $M' \in \alpha'$, в которой прямая OM пересекает плоскость α'. Предполагается, что плоскости α, α' не проходят через центр O (рис. 64).

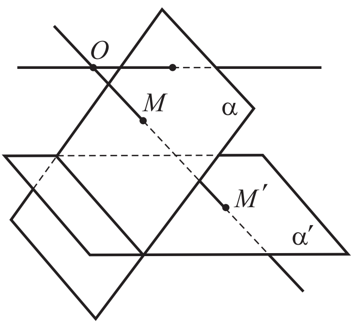


Рис. 64

Свойства плоских фигур, сохраняющиеся при проектировании, называются *проективными*, учение о них и составляет предмет *проективной геометрии на плоскости.*

Однако тут есть существенная трудность. Она состоит в том, что если плоскость α' не параллельна плоскости α, то прямая OM может оказаться параллельной плоскости α'. Тогда точки M' не существует: точка M не имеет проекции. Чтобы обойти это, считают, что такая прямая OM «пересекает» плоскость α', но только в «бесконечно удаленной точке». Когда точка M, двигаясь по прямой в плоскости α, приближается к такому положению, то прямая OM становится параллельной плоскости α', проекция точки M — точка M' — удаляется в бесконечность. Но теперь мы считаем, что она приближается к соответствующей бесконечно удаленной точке (как на рис. 68).

Считается, что все бесконечно удаленные точки плоскости образуют на ней одну «бесконечно удаленную прямую». Это обосновывается из свойств проектирования. Для этого рассмотрим проекции прямых и докажем, что проекция прямой, вообще говоря, представляет собой прямую.

Действительно. Возьмем в плоскости α какую-нибудь прямую a. Через нее и через центр проекции O проходит плоскость (рис. 65). Эта плоскость β содержит все прямые, вдоль которых проектируются точки прямой a. Поэтому проекции всех этих точек, образуя проекцию

прямой a, образуют вместе с тем пересечение плоскости β с плоскостью проекции α'. А пересечение двух плоскостей представляет собою прямую a'. Значит, проекция прямой есть прямая. Кроме одного случая!

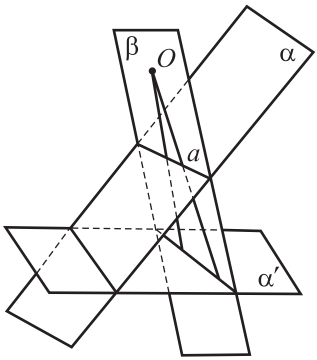


Рис. 65

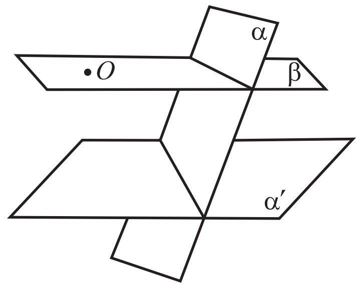


Рис. 66

Если плоскость β параллельна плоскости α', то никакой такой прямой a' нет.

Но прямые, проходящие через центр O и лежащие в плоскости β, и тем самым параллельные плоскости a', по условию «пересекают» ее в бесконечно удаленных точках. Соответственно, мы считаем, что плоскость β «пересекает» плоскость α' по «бесконечно удаленной прямой», которая образуется всеми бесконечно удаленными точками плоскости α' (рис. 66).

Кроме того, можно заметить, что у всех прямых плоскости α, пересекающих прямую a, точки их пересечения с a проектируются в бесконечно удаленные точки. Поэтому считать, что проекция прямой есть прямая, можно, только учитывая бесконечно удаленные точки.

Именно имея в виду все эти оговорки, мы и сказали, что проекция прямой, «вообще говоря», представляет собой прямую. Но теперь можно утверждать:

Проекция всякой прямой представляет собой прямую, поскольку к прямым присоединяется бесконечно удаленная прямая, образованная всеми бесконечно удаленными точками плоскости.

На плоскости α тоже вводится бесконечно удаленная прямая, так

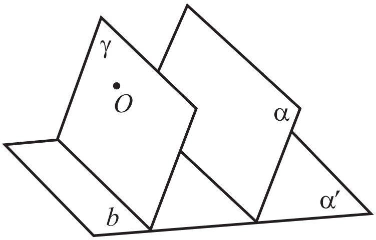


Рис. 67

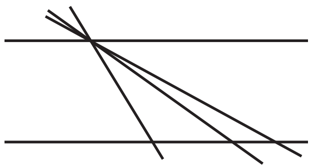


Рис. 68

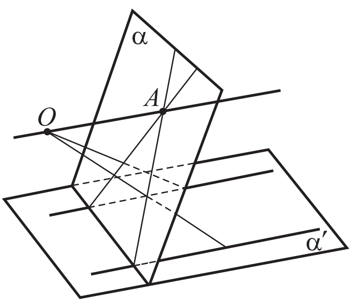


Рис. 69

что и ей приписывается проекция. Это та прямая b, по которой плоскость α' пересекается плоскостью γ, проходящей через центр O параллельно плоскости α (рис. 67; плоскость γ и «пересекает» α по бесконечно удаленной прямой).

На всякой прямой есть одна бесконечно удаленная точка, в которой она пересекает бесконечно удаленную прямую. Параллельные прямые сходятся — «пересекаются» в бесконечно удаленной точке. Это соответствует тому, что когда точка, двигаясь по прямой, удаляется в бесконечность, пересекающиеся в ней прямые «приближаются к параллельности» (рис. 68).

То же получается при проектировании. Пусть A — такая точка плоскости α, что прямая OA параллельна плоскости проекции α', так что точке A соответствует бесконечно удаленная точка A'. Прямые, проходящие на плоскости α через точку A, в проекции уже не пересекаются, так как у точки A нет проекции, т. е. в проекции получаются параллельные прямые. Но так как точке A' сопоставлена бесконечно удаленная точка A', то мы должны считать, что эти параллельные прямые пересекаются в точке A' (рис. 69).

Итак, плоскость дополняется *бесконечно удаленными точками;* они образуют *бесконечно удаленную прямую;* каждая другая (обычная) прямая содержит одну бесконечно удаленную точку, и все параллельные друг другу прямые пересекаются в одной бесконечно удаленной точке. Тем самым *всякие две прямые пересекаются*: параллельные — в бесконечно удаленной точке, а бесконечно удаленная прямая пересекается со всякой прямой в ее бесконечно удаленной точке.

При этих условиях (уже без всяких оговорок) *прямые при проек-*

тировании отображаются на прямые, и точки, лежащие на одной прямой, отображаются в точки, лежащие на одной прямой.

Прямая, дополненная бесконечно удаленной точкой, называется ***проективной прямой.*** У прямой бесконечно удаленная точка одна, и когда точка движется по прямой в бесконечность в одну или другую сторону, она в обоих направлениях «стремится» к бесконечно удаленной точке. Это значит, что прямая замыкается в бесконечно удаленной точке. Проективная прямая представляет собой замкнутую линию.

Плоскость, дополненная бесконечно удаленными точками и вместе с ними — бесконечно удаленной прямой, называется ***проективной плоскостью*** при условии, что эти точки признаются равноправными с обычными точками и бесконечно удаленная прямая — равноправной с обычными прямыми. Равноправность эта проявляется в том, что все точки и прямые переходят одни в другие при проектировании. Конечно, это понятие «равноправности» несколько неопределенно; строгое его определение будет дано дальше.

Спроектировав плоскость α на α', можно затем спроектировать α' на α из другого центра или параллельным проектированием. В результате получается проективное отображение проективной плоскости α самой на себя. Подобную операцию можно повторять. При этом как при каждом проектировании, так и при их композиции, прямые будут отображаться на прямые.

Вообще *проективным преобразованием* проективной плоскости называется ее отображение самой на себя, при котором прямые отображаются на прямые. (Преобразуемость одних прямых в другие и определяет их «равноправие».)

Проективным называется свойство, сохраняющееся при любых проективных преобразованиях.

Проективная геометрия на плоскости изучает проективные свойства фигур на проективной плоскости, не выделяя бесконечно удаленные элементы; для нее все точки и все прямые одинаковы. Фигуры, преобразуемые одна в другую, имеют одни и те же проективные свойства; они с точки зрения проективной геометрии эквивалентны.

Переход от проективной плоскости к аффинной. Теорема Дезарга. Дополнение обычной плоскости бесконечно удаленной прямой позволяет в целом ряде случаев не отличать параллельные прямые от пересекающихся и этим придает единообразие формулировкам и доказательствам теорем.

С другой стороны, поскольку проективная плоскость получается

из обычной прибавлением прямой, то, приняв какую-либо прямую на проективной плоскости за бесконечно удаленную, получаем обычную аффинную плоскость. Тогда теорема проективной геометрии сводится к теореме аффинной геометрии, и ее можно доказывать, пользуясь аффинной и даже евклидовой геометрией.

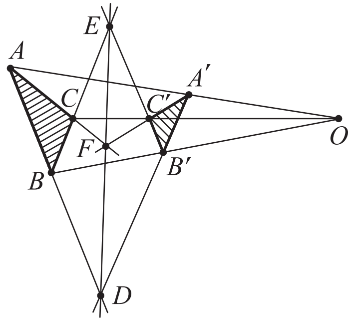


Рис. 70

Пример, демонстрирующий пользу обоих сделанных замечаний, представляет теорема Дезарга. Формулируем ее для проективной плоскости (рис. 70).

Теорема 1 (Дезарг). *Пусть у двух треугольников вершины и, соответственно, стороны приведены в соответствие. Тогда если при этом оказывается, что прямые, проходящие через соответственные вершины, пересекаются в одной точке, то точки пересечения прямых, проходящих вдоль соответственных сторон, лежат на одной прямой.*

И обратно: *если точки пересечения прямых, проходящих вдоль соответственных сторон, лежат на одной прямой, то прямые, проходящие через соответственные вершины, пересекаются в одной точке.*

Обозначим, как обычно, соответствующие вершины треугольников A, B, C и A', B', C'. Точки пересечения прямых AB и $A'B'$, BC и $B'C'$, CA и $C'A'$ обозначим D, E, F. Тогда теорема выглядит так.

Если прямые AA', BB', CC' пересекаются в одной точке, то точки D, E, F лежат на одной прямой, и обратно: если D, E, F — на одной прямой, то AA', BB', CC' пересекаются в одной точке.

Заметим, что в проективной геометрии прямая — замкнутая линия, поэтому отрезок не определяется своими концами: есть два отрезка с одними и теми же концами. Соответственно, треугольник ABC не является определенным в обычном смысле. Поэтому, может быть, лучше говорить о двух тройках точек A, B, C и A', B', C', не лежащих каждая на одной прямой.

Переведем теперь теорему Дезарга на язык евклидовой геометрии. Тогда для прямых, проходящих через соответственные вершины, надо различать два случая: 1) либо они пересекаются, 2) либо они параллельны. Для прямых, проходящих вдоль соответственных сторон, придется различать три случая: 1) они пересекаются в точках од-

ной прямой, 2) они параллельны (три пары параллельных прямых), 3) прямые одной пары параллельны, прямые двух других пар пересекаются в точках, лежащих на прямой, параллельной прямым первой пары.

Итого, в теореме при ее формулировке для обычной плоскости будет $2 \times 3 = 6$ случаев.

С другой стороны, можно превратить данную теорему в теорему на обычной плоскости, приняв прямую, на которой лежат точки пересечения прямых, идущих вдоль сторон, за бесконечно удаленную. Тогда теорема примет такой вид.

Теорема 2. *Пусть вершины и соответственно стороны двух треугольников приведены в соответствие.*

Пусть при этом оказывается, что 1) *прямые, соединяющие соответственные вершины, пересекаются в одной точке или параллельны,* 2) *соответственные стороны в двух парах параллельны. Тогда и стороны третьей пары параллельны.*

Обратно: если соответственные стороны параллельны (*в каждой из трех пар*), *то прямые, проходящие через соответственные вершины, пересекаются либо параллельны.*

В этом виде обе части теоремы, прямая и обратная, вполне доступны доказательству на уровне школьного курса. Докажем, например, вторую часть.

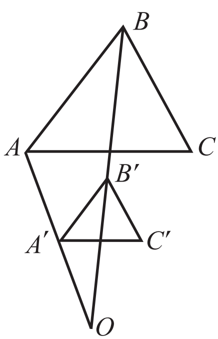


Рис. 71

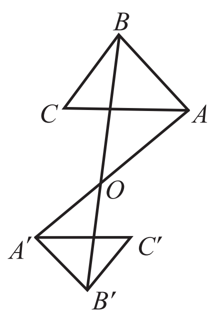


Рис. 72

Пусть $AB||A'B'$, $BC||B'C'$, $CA||C'A'$. Проведем прямые AA', BB'. Допустим, они пересекаются в некоторой точке O.

Так как $AB||A'B'$, то

$$OA/OA' = OB/OB'. \qquad (1)$$

Тут возможны два случая: 1) точки A, A' как и B, B' лежат с одной стороны от O (рис. 71), 2) они с разных сторон от O (рис. 72). Произведем гомотетию с центром O, которая переведет точку A' в A. Если эти точки с разных сторон от O, то коэффициент гомотетии отрицательный.

Ввиду пропорциональности отрезков (1) точка B' перейдет в B. Прямые $A'C'$ и $B'C'$, как параллельные прямым AC, BC, перейдут в эти прямые. Вместе с этим точка их пересечения C' переходит в C.

Следовательно, эти точки C, C' лежали на одной прямой, проходящей через O.

Если $AA'||BB'$, то тот же результат дает параллельный перенос, совмещающий A с A'. (Убедитесь!)

Итак, вторая часть теоремы доказана.

Первая ее часть доказывается сходным приемом, и читатель сам проделает это доказательство. □

Перейти от доказываемой таким образом теоремы из элементарной геометрии к теореме Дезарга в ее общем виде можно совсем просто и не ссылаясь на проективную геометрию.

Пусть на плоскости α даны треугольники ABC, $A'B'C'$. Воспользуемся введенными выше обозначениями (рис. 70). Пусть D, E — точки пересечения прямых AB, $A'B'$ и BC, $B'C'$. (Если, скажем, $BC||B'C'$, то берем прямые CA, $C'A'$, если же $CA||C'A'$ или $AB||A'B'$, то это случай, который уже рассмотрен в доказанной теореме.) Проводим прямую DE. Берем точку O — центр проекции вне плоскости α и проводим плоскость β через O и прямую DE. Проводим плоскость $\gamma||\beta$ и проектируем на нее плоскость α. Так как $\gamma||\beta$, то прямая DE спроектируется в бесконечно удаленную прямую, т. е., попросту говоря, прямые AB, $A'B'$ и BC, $B'C'$ спроектируются в параллельные. Мы получим, таким образом, конфигурацию, рассматриваемую в теореме 2. Эта теорема доказана, а значит, — возвращаясь на плоскость α, — доказана и теорема Дезарга. □

Содержащийся в изложении прием свед́ения теорем и построений проективной геометрии к элементарной геометрии продуктивно работает и в других случаях. Рассмотрим несколько примеров в качестве задач.

Теорема 3 (Паскаль). *Прямые, проходящие вдоль противоположных сторон шестиугольника, вписанного в коническое сечение, пересекаются в точках, лежащих на одной прямой.*

Переведите в теорему элементарной геометрии и докажите.

Теорема 4. *Если точка A лежит вне конического сечения, то через нее проходят две касательные к нему. Пусть $p(A)$ — прямая, проходящая через точки касания. Теорема утверждает: для точек A, лежащих на прямой, не пересекающей коническое сечение, прямые $p(A)$ проходят через одну точку.*

Исторические замечания. Изучение центрального проектирования возникло в связи с потребностью правильного изображения предметов, как мы их видим, — в перспективе, можно сказать, в проекции на плоскость зрения (упрощенно можно считать, что изображение предметов на сетчатке глаза получается проектированием по лучам света с центром проекции в центре зрачка). Перспективным (линейной перспективой) и называется способ изображения пространственных фигур на плоскости посредством центрального проектирования. Глядя вдаль, мы видим параллельные линии сходящимися в общую точку: плоскость равнины, простирающейся перед взором, проектируется на плоскость зрения.

Теория перспективы возникла из потребностей архитектуры и живописи. Некоторые ее законы были известны еще древнегреческим геометрам. Ею занимались крупнейшие художники Возрождения Леонардо да Винчи (1452–1519) и Альбрехт Дюрер (1471–1521); ими были написаны книги о перспективе.

Основы проективной геометрии заложил французский инженер, архитектор и математик Жерар Дезарг (1591–1661). Ему принадлежит принципиальная заслуга — понятие о бесконечно удаленных точках, он же открыл и доказал одну из основных теорем проективной геометрии — теорему Дезарга, которую мы только что доказали.

Проективная геометрия была развита и оформилась как особая область геометрии в работе французского инженера и геометра Жана Виктора Понселе (1788–1867) «Трактат о проективных свойствах фигур», вышедшей в 1822 г. Понселе был офицером наполеоновской армии, попал в плен и написал свой трактат в 1813–1814 гг., находясь в плену в Саратове.

Достойно внимания, что эта область геометрии не только возникла из потребностей практики, но и была обязана своим развитием художникам, архитекторам и инженерам. Только позже ей было дано самостоятельное аксиоматическое обоснование.

§ 2. Проективная плоскость как связка прямых. Координаты

Когда плоскость дополнена бесконечно удаленными точками, то между ее точками и прямыми, проходящими через центр проекции O, устанавливается взаимно однозначное соответствие. Прямые, параллельные плоскости, тоже «пересекают» ее, но в бесконечно удаленных точках.

Множество всех прямых, проходящих через одну точку O, называется ***связкой прямых*** *с центром* O. Установленное взаимно однозначное соответствие точек дополненной плоскости и прямых связки позволяет рассматривать вместо точек — прямые связки, вместо дополненной плоскости — связку прямых. (Выгода состоит в том, что в связке все прямые равноправны, на плоскости же нужно еще условиться считать бесконечно удаленные точки равноправными с обычными и определить, в каком смысле они считаются равноправными.)

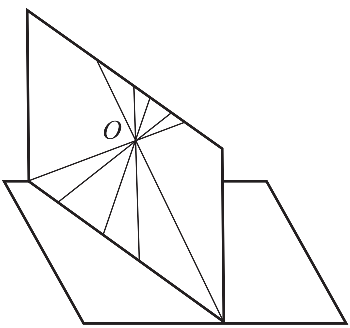


Рис. 73

Прямые, проходящие через центр O и точки прямой на проектируемой плоскости (включая бесконечно удаленную точку), покрывают плоскость, проходящую через центр O (рис. 73). Поэтому если мы вместо точек дополненной плоскости рассматриваем прямые связки с центром O, то роль прямых будут играть плоскости, проходящие через O, — «плоскости связки».

Так мы приходим к определению. ***Проективной плоскостью*** называется связка прямых в евклидовом (или аффинном) пространстве при условии, что прямые связки считаются точками, а плоскости связки — прямыми.

При таком взгляде строение самих прямых не играет роли. Выступая как точки, они становятся тем, что, по определению Евклида, «не имеет частей». Каждая из них фигурирует как нечто целое. Такой взгляд служит прекрасной иллюстрацией отвлеченного понимания геометрии, когда она относится к объектам «произвольной природы». В данном случае роль точек играют прямые связки в евклидовом пространстве.

Теперь мы можем по-новому определить, что понимается

под проективными преобразованиями и проективной геометрией.

Проективной геометрией на проективной плоскости называется аффинная геометрия связки прямых. При этом прямые связки рассматриваются как точки этой геометрии без всякого внимания к их внутренней структуре.

Проективным преобразованием проективной плоскости называется аффинное преобразование связки (сохраняющее ее центр), точнее: отображение множества прямых связки на себя, происходящее при каком-либо аффинном ее преобразовании. (Это определение равносильно данному выше, гласящему, что проективные преобразования — те, которые отображают прямые на прямые, т. е. плоскости связки — на плоскости связки. Докажите!)

Хотя представление проективной плоскости в виде связки прямых позволяет дать точные определения без особых оговорок о бесконечно удаленной прямой, все же замена плоскости связкой делает наглядные представления более сложными. Поэтому рассуждают обычно на плоскости и рисуют соответствующие рисунки. Однако каждый способ рассмотрения имеет свои преимущества, и ими можно пользоваться совместно. В следующем пункте мы рассмотрим с этой точки зрения конические сечения.

Конические сечения. Может быть, самый яркий пример, поясняющий проективную точку зрения, по которой фигуры, проективно преобразуемые друг в друга, представляют не более как разные экземпляры одной и той же фигуры, дают конические сечения. С проективной точки зрения между ними мало разницы. Разница между ними обусловлена только тем, как они располагаются относительно бесконечно удаленной прямой. Эллипс не имеет с нею ничего общего, параболы она касается, а гиперболу она пересекает. Но в проективном смысле особой бесконечно удаленной прямой просто нет. Ведь все прямые равноправны. Поэтому не различаются и конические сечения.

Раскроем эти краткие замечания более точно. Воспользуемся представлением проективной плоскости связкой прямых.

В связке прямых с центром O круговой конус с вершиной O, вообще конус второго порядка с вершиной O, представляет собой вполне определенную фигуру, а его сечения разными плоскостями — это конические сечения. В чем их различие?

Пусть коническое сечение K получено в пересечении конуса C с плоскостью α (рис. 74). Плоскость β, проходящая через центр O параллельно α, представляет в связке ту прямую, которая оказывается бесконечно удаленной на плоскости α (прямые, лежащие в плоскости

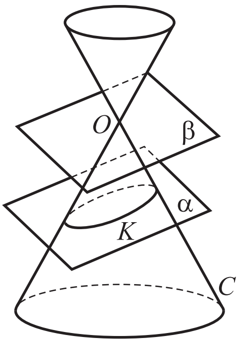


Рис. 74

β, параллельны плоскости α). Следовательно, конические сечения различаются тем, какая плоскость связки играет в ней роль бесконечно удаленной прямой.

Итак, конические сечения — КВП, не вырождающиеся и не распадающиеся на прямые, — представляются в связке конусами второго порядка. Все эти конусы преобразуются один в другой аффинными преобразованиями связки. Поэтому с точки зрения проективной геометрии они эквивалентны — «проективно равны». Никаких конических сечений разных типов в проективной геометрии нет. Разные конические сечения есть на обычной (евклидовой или аффинной) плоскости. Эта плоскость возникает из проективной, когда выделяется прямая, которая считается бесконечно удаленной. В связке это соответствует тому, что в ней выделяется плоскость связки, представляющая в связке выделяемую прямую. В зависимости от расположения этой плоскости относительно конуса и получаются разные случаи соответственно разным коническим сечениям. Случай, когда плоскость не имеет с конусом общих точек, кроме вершины, соответствует эллипсу; если плоскость касается конуса — это соответствует параболе, а когда плоскость пересекает конус — гиперболе.

Когда в евклидовом пространстве конус пересекается разными плоскостями, получаются, скажем, разные эллипсы. Но все эллипсы аффинно эквивалентны. Их различие появляется только в евклидовой — метрической геометрии.

То же верно для других типов конических сечений.

Однородные координаты. Введем в пространстве аффинные координаты x_1, x_2, x_3 с началом в центре связки O. Если a — прямая связки и $A(a_1,\ a_2,\ a_3)$ — какая-либо ее точка, отличная от O, то координаты любой точки $X(x_1,\ x_2,\ x_3)$, отличной от O, на прямой a представляются в виде

$$x_1 = ta_1, \quad x_2 = ta_2, \quad x_3 = ta_3, \quad t \neq 0.$$

Таким образом, каждая прямая связки задается координатами $x_1,\ x_2,\ x_3$, не равными одновременно нулю и определенными с точностью до произвольного общего множителя, отличного от нуля. Это выражают, говоря, что прямую задает отношение этих координат:

$(x_1 : x_2 : x_3)$. Так определяемые с точностью до множителя координаты называются *однородными.* И так как прямая связки — это точка проективной плоскости, то эти же однородные координаты служат координатами точек проективной плоскости.

Уравнение плоскости, проходящей через начало, имеет вид

$$a_1x_1 + a_2x_2 + a_3x_3 = 0 \quad (a_1^2 + a_2^2 + a_3^2 \neq 0). \tag{1}$$

Это же представляет уравнение прямой на проективной плоскости в однородных координатах.

Поскольку однородные координаты определены с точностью до множителя, то соотношение между ними может иметь смысл только, если оно само однородно, т. е. сохраняется при умножении входящих в него координат на любой множитель, отличный от нуля. Например, равенство $x_1 + x_2 = 1$ в однородных координатах бессмысленно, так как не сохраняется при умножении x_1, x_2 на какое-либо $t \neq 0$ $(\neq 1)$.

Связь с обычными координатами на плоскости. Выясним связь однородных координат с обычными аффинными или прямоугольными координатами на плоскости.

Рассмотрим связку прямых с центром O. Возьмем какую-нибудь плоскость α, не проходящую через O, и введем на ней координаты x, y. (Ради простоты можно представлять их прямоугольными, а начало их O_α — расположенным в основании перпендикуляра, опущенного из O на плоскость α (рис. 75).) Введем в пространстве координаты x_1, x_2, x_3 с началом в центре связки O. Ось x_3 направим из O через O_α, а оси x_1, x_2 проведем параллельно осям x, y. Тогда плоскость $x_3 = 0$ параллельна плоскости α, и мы можем считать, что плоскость α представляется уравнением $x_3 = 1$. В плоскости $x_3 = 0$ координаты x_1, x_2 вводим так же, как x, y — на плоскости α, т. е. переносим их значения по прямым, параллельным оси x_3. Тогда точки плоскости α имеют координаты $x_1 = x$, $x_2 = y$, $x_3 = 1$.

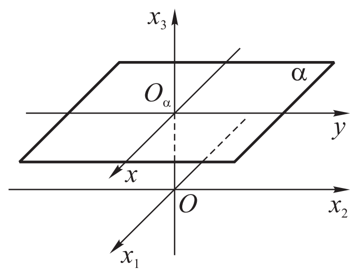


Рис. 75

Теперь перейдем к однородным координатам, т. е. введем произвольный множитель $t \neq 0$, так что для точек плоскости будет

$$x_1 = tx,\ x_2 = ty,\ x_3 = t \quad (t \neq 0). \tag{2}$$

Отсюда
$$x = \frac{x_1}{x_3}, \quad y = \frac{x_2}{x_3}. \tag{3}$$

Прямые связки, параллельные плоскости α, «пересекают» ее в бесконечно удаленных точках. Они лежат в плоскости $x_3 = 0$. Поэтому $x_3 = 0$ представляет уравнение той прямой, которая на плоскости α оказывается бесконечно удаленной.

Это и соответствует тому, что по формулам (3) при $x_3 = 0$, грубо говоря, координаты x и y становятся бесконечными. Или, выражаясь более корректно, значение x_3 недопустимо, а это и значит, что на плоскости α нет бесконечно удаленной прямой. Она присоединяется к α введением однородных координат с допущением $x_3 = 0$.

Формулы (3) задают переход от аффинных координат x, y к однородным. Достаточно в формулы с x, y подставить вместо них x_1/x_3, x_2/x_3 и для удобства избавиться от знаменателя. Эти же формулы задают переход от однородных координат x_1, x_2, x_3 к аффинным x, y (когда исключается значение $x_3 = 0$, соответствующее бесконечно удаленной плоскости).

Теорема 5. *Проективное преобразование проективной плоскости представляется в однородных координатах линейным (однородным) преобразованием вида*

$$\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x'_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{aligned} \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0. \tag{4}$$

Доказательство. Проективное преобразование проективной плоскости, когда она представлена как связка прямых, является, по определению, аффинным преобразованием связки, сохраняющим ее центр. Если начало аффинных координат положено в центре связки, то аффинные преобразования представляются в координатах формулами вида (4). Аффинные координаты с началом в центре связки являются однородными координатами на проективной плоскости (представленной связкой прямых). Этим теорема 5 доказана. □

Теорема 6. *Проективное преобразование плоскости представляется в аффинных координатах как дробно-линейное, т. е.*

$$\begin{aligned} x' &= \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}, \\ y' &= \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}, \end{aligned} \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0.$$

Докажите!

Проективное аффинное отображение области. Если рассматривать отображение ограниченных областей на ограниченные области, так что никакая бесконечно удаленная точка не вмешивается, то проективное отображение можно определить очень просто.

Проективное отображение ограниченной области — это такое ее взаимно однозначное отображение, при котором отрезки отображаются на отрезки.

Аффинное отображение, напомним, подчиняется дополнительному требованию, что параллельные отрезки отображаются на параллельные (теорема III.2.5).

Доказательство мы не приводим.

КВП на проективной плоскости. Кривые второго порядка задаются в однородных координатах однородными уравнениями второй степени (это ясно по свойству однородных координат и из перехода в уравнении КВП от прямоугольных координат к однородным координатам); т. е. в таком уравнении слева стоит квадратичная форма:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{23}x_2x_3 + 2a_{31}x_3x_1 = 0.$$

По теореме, доказанной в § 2 гл. V ч. 1, квадратичную форму можно линейным преобразованием привести к «сумме квадратов», так что уравнение примет вид:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 = 0.$$

Произведем еще линейное преобразование

$$x_1' = \sqrt{|a_{11}|}x_1, \quad x_2' = \sqrt{|a_{22}|}x_2, \quad x_3' = \sqrt{|a_{33}|}x_3.$$

Тогда уравнение примет вид

$$e_1x_1'^2 + e_2x_2'^2 + e_3x_3'^2 = 0,$$

где e_1, e_2, e_3 равны $+1$ или -1, или 0, в зависимости от знака a_{11}, a_{22}, a_{33}. Умножение на -1 не нарушает уравнения, поэтому можно считать, что число положительных квадратов не меньше числа отрицательных и, разумеется, переменные можно обозначать произвольно. В результате получаем пять возможных случаев:

$$1)\ x_1^2 + x_2^2 + x_3^2 = 0, \quad 2)\ x_1^2 + x_2^2 - x_3^2 = 0,$$
$$3)\ x_1^2 + x_2^2 = 0, \quad 4)\ x_1^2 - x_2^2 = 0, \quad 5)\ x_1^2 = 0.$$

Первое уравнение представляет пустое множество, т. к. все однородные координаты не могут равняться нулю. Второе представляет коническое сечение — конус в связке прямых.

Третье уравнение, $x_1^2 + x_2^2 = 0$, представляет в связке одну прямую — ось x_3, т. е. точку $(0, 0, x_3)$ на проективной плоскости.

Следующее уравнение 4) представляет на проективной плоскости пару пересекающихся прямых

$$x_1 + x_2 = 0, \quad x_1 - x_2 = 0.$$

Последнее уравнение 5) представляет в связке плоскость $x_1 = 0$, т. е. прямую на проективной плоскости.

Те же случаи получаются из выводов § 7 гл. II ч. 1, где рассматривались КВП на евклидовой плоскости: три конических сечения теперь объединены, также объединены пересекающиеся и параллельные прямые; параллельные пересекаются в бесконечно удаленной точке.

Замечание. Однако тут есть особенность. Рассмотрим второе уравнение, оно представляет коническое сечение. В нем значение $x_3 = 0$ невозможно, так как если $x_3 = 0$, то $x_1^2 + x_2^2 = 0$, и должно быть $x_1 = x_2 = 0$, т. е. все три координаты — нули, что исключено. Итак, $x_3 \neq 0$, а потому можно разделить на x_3 и получить из формулы 2)

$$x^2 + y^2 = 1.$$

Это окружность или эллипс в подходящих аффинных координатах. А где гипербола с параболой? Разделив на x_2 и положив $x = x_3/x_2$, $y = x_1/x_2$, получим:

$$x^2 - y^2 = 1, \quad \text{т. е.} \quad y^2 + 1 - x^2 = 0.$$

Это гипербола. Но исключено значение $x_2 = 0$, которое в самом уравнении $x_1^2 + x_2^2 - x_3^2 = 0$ возможно. Исключая значение $x_2 = 0$, мы делаем прямую $x_2 = 0$ бесконечно удаленной, а так как значение $x_2 = 0$ возможно, то кривая пересекает бесконечно удаленную прямую. А где же все-таки парабола? Положим

$$x_3 - x_2 = \overline{x}_2, \quad x_3 + x_2 = \overline{x}_3.$$

Тогда уравнение 2) станет

$$x_1^2 - \overline{x}_2\overline{x}_3 = 0.$$

Или, полагая $y = x_1/\overline{x}_3$, $x = \overline{x}_2/\overline{x}_3$,

$$y^2 = x.$$

Заметим еще, что в § 1 гл. I, ч. 1, рассматривая КВП на евклидовой плоскости, мы получили пустое множество при двух уравнениях

$$x^2 + y^2 + 1 = 0, \quad x^2 + 1 = 0.$$

Первое соответствует уравнению 1). А второе?

В однородных координатах, полагая $x = x_1/x_3$, получим $x_1^2 + x_3^2 = 0$. Это точка $(0,\ x_2,\ 0)$, а не пустое множество! Однако эта точка лежит на бесконечно удаленной прямой $x_3 = 0$. Поэтому ее на евклидовой (и на аффинной) плоскости нет.

Обычная точка получается сжатием окружности (или эллипса) к центру. Точка на бесконечно удаленной прямой получается, если «раздвигать» гиперболу, беря ее вещественную полуось $a \to \infty$ и мнимую полуось $b = \text{const}$, например $b = 1$. Это можно понимать как сжатие к бесконечно удаленной точке.

§ 3. Принцип двойственности

В связке между прямыми и плоскостями имеется простое взаимно однозначное соответствие: каждой прямой соответствует перпендикулярная ей плоскость, и каждой плоскости — перпендикулярная ей прямая. Другими словами, можно сказать: в связке имеется взаимно однозначное соответствие прямых и плоскостей по их взаимной перпендикулярности (рис. 76).

Если прямая a лежит в плоскости β, то перпендикулярная ей плоскость α проходит через прямую b, перпендикулярную плоскости β. Это значит, что если a и α соответствуют друг другу по перпендикулярности и a содержится в плоскости β, то α проходит через прямую b, соответствующую плоскости β (рис. 77). Таким образом, указанное сопоставление прямых и плоскостей устанавливает соответствие между тем, что прямая лежит в плоскости, и тем, что плоскость проходит через прямую.

Для того чтобы выразить это соответствие более четко, вводят термин «инцидентность». Прямая *инцидентна* плоскости, если лежит в ней; плоскость *инцидентна* прямой, если содержит эту прямую. Таким образом, они друг другу инцидентны. Отношение инцидентности взаимно.

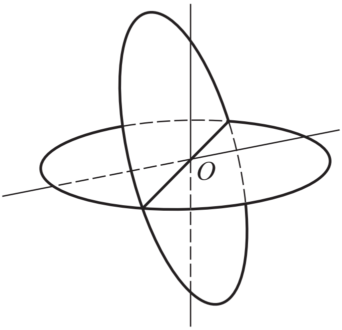


Рис. 76

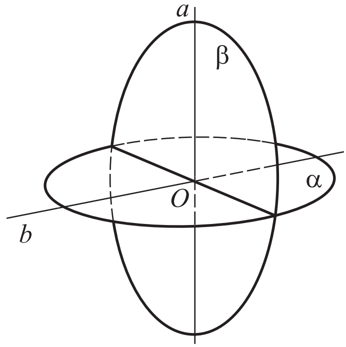


Рис. 77

Это позволяет выразить отмеченное свойство соответствия прямых и плоскостей очень просто: *соответствие прямых и плоскостей по перпендикулярности сохраняет инцидентность.*

Если от связки перейти к проективной плоскости — если толковать связку как проективную плоскость, то прямые связки представят точки, а плоскости связки — прямые. Соответствие прямых и плоскостей связки будет соответствием точек и прямых проективной плоскости. Инцидентность будет инцидентностью точек и прямых, и она сохраняется при данном соответствии: точкам, лежащим на прямой, соответствуют прямые, проходящие через точку, и обратно. Такое соответствие называют *коррелятивным*, или *полярным*, — ***полярностью***. Прямая, соответствующая точке, — это ее *поляра*; точка, соответствующая прямой — ее *полюс*. При этом можно иметь в виду любое соответствие точек и прямых, сохраняющее инцидентность, не обязательно устанавливаемое так, как это мы сделали, сопоставляя в связке прямые и плоскости по взаимной перпендикулярности. Мы только указали конкретный вид полярного соответствия и, что самое главное, доказали этим, что такое соответствие существует. (Кстати, на обычной плоскости его нет!)

Существование полярного соответствия приводит к замечательному выводу, называемому ***принципом двойственности.*** *Всякому верному утверждению о точках и прямых проективной плоскости, касающемуся их инциденции, соответствует также верное утверждение, в котором точки и прямые меняются местами с сохранением инциденций.* Короче: каждой теореме отвечает двойственная ей по полярности теорема. Доказав одну, мы «даром» получаем вместе с нею и другую — ей двойственную.

Пример тому дает теорема Дезарга. Сформулируем ее — первую из двух доказанных выше, см. теорему 1, § 1.

Пусть даны две тройки точек A, B, C и A', B', C', не лежащих каждая на одной прямой — не инцидентных каждая с одной прямой, — причем точки троек поставлены в соответствие: A с A' и т. д. Пусть прямые AA', BB', CC' пересекаются в одной точке, иначе говоря, прямые, инцидентные с точками A и A', B и B', C и C', инцидентны с одной точкой P. Тогда, утверждает теорема, точки пересечения трех прямых: AB с $A'B'$, BC с $B'C'$, CA с $C'A'$ лежат на одной прямой. Иначе говоря, три точки, инцидентные с прямыми AB, $A'B'$ и т. д., инцидентны с одной прямой.

Сформулируем двойственное утверждение.

Пусть даны две тройки прямых a, b, c и a', b', c', не инцидентные каждая с одной точкой (т. е. не проходящие каждая через одну точку), причем прямые троек поставлены в соответствие: a с a' и т. д. Пусть точки пересечения прямых a с a', b с b', c с c' лежат на одной прямой (инцидентны с одной прямой p). Тогда три прямые, проходящие через точки пересечения прямых a с b и a' и b' и т. д., проходят через одну точку — иначе говоря, эти три прямые инцидентны с одной точкой.

Читатель узнает в этой формулировке теорему, обратную предыдущей, — вторую часть теоремы Дезарга. По принципу двойственности она уже доказана, когда доказана предыдущая теорема. Доказывая ее, мы «зря старались», не зная еще принципа двойственности.

Общее полярное соответствие. Мы определили полярность, отправляясь от соответствия прямых и плоскостей связки по взаимной перпендикулярности. Но это можно обобщить, применяя к данной полярности любое проективное преобразование. Действительно, проективное преобразование сохраняет инцидентность, переводя точки в точки и прямые в прямые. Если его добавить к полярному преобразованию, то получится преобразование, опять сохраняющее инцидентность и сопоставляющее точкам прямые, а прямым — точки, т. е. получается полярность.

В связке это соответствует тому, что к сопоставлению прямых и плоскостей по взаимной перпендикулярности применяется аффинное преобразование.

Дополнение к принципу двойственности. Соответствие между прямыми и плоскостями связки по их взаимной перпендикулярности очевидным образом взаимно непрерывно. (Просто угол между плоскостями равен углу между перпендикулярными им прямыми, так

что когда один из них стремится к нулю, то и другой стремится к нулю.)

Когда прямая связки описывает плоский угол между двумя прямыми a, b, то перпендикулярная ей плоскость зачерчивает двугранный угол между соответствующими плоскостями α, β (а вернее, два взаимно вертикальных двугранных угла).

На проективной плоскости плоскому углу отвечает отрезок с концами в точках A, B, соответствующих прямым a, b. Двугранному углу отвечает угол между прямыми — часть плоскости, зачерчиваемая прямой, поворачивающейся вокруг данной точки от одного положения к другому.

Двойственность гладких кривых. Гладкая кривая имеет в каждой точке касательную прямую и является огибающей семейства своих касательных. Ее можно рассматривать двояко: как определяемую своими точками или как определяемую прямыми. Касательная — это прямая, «проходящая через две бесконечно близкие точки кривой», а точка огибающей семейства прямых — это точка «пересечения двух бесконечно близких прямых семейства». Более точно, касательная к кривой в точке A — это прямая, служащая пределом прямых, проходящих через точку A и другую точку кривой — M, когда M «стремится» к A. Точка огибающей кривой на прямой a семейства — это точка, служащая пределом точек, лежащих на прямой a и другой прямой m семейства, когда m стремится к a. В этих определениях ясно видна их двойственность, так что по принципу двойственности можно переходить от одного взгляда на кривые к другому — двойственному. Это вовсе не значит, что любая гладкая кривая сама себе двойственна в том смысле, что полярное преобразование переводит ее точки в касательные, и обратно. Но это значит, что полярность сопоставляет выпуклой кривой выпуклую кривую, меняя ролями точки и касательные.

Простейший пример двойственности для кривых представляют вписанные и описанные многоугольники. Вписанный многоугольник задается вершинами — точками, лежащими на кривой; описанный многоугольник задается сторонами — прямыми, касающимися двойственной кривой. По двойственности сторонам и диагоналям вписанного многоугольника соответствуют точки пересечения сторон описанного многоугольника, и обратно.

Рассмотрим для примера теорему Паскаля. *У шестиугольника, вписанного в коническое сечение, точки пересечения прямых, проходящих вдоль противоположных сторон, лежат на одной прямой.*

Двойственной ей будет такая теорема. *У шестиугольника, описан-*

ного вокруг кривой, двойственной коническому сечению, диагонали, соединяющие противоположные вершины, пересекаются в одной точке. (Читатель проверит двойственность этого утверждения теореме Паскаля IV.1.3.)

Оказывается, как мы сейчас докажем, что кривая, двойственная коническому сечению, сама является коническим сечением. Теорема, двойственная теореме Паскаля, называется теоремой Брианшона (по имени доказавшего ее математика). Ее достаточно формулировать для окружности (поскольку все конические сечения в окружность проективно преобразуются):

У шестиугольника, описанного около окружности, диагонали, соединяющие противоположные вершины, пересекаются в одной точке.

Автополярность конических сечений. Конические сечения не только двойственны друг другу, но более того:

Теорема 1. *Каждое коническое сечение полярно самому себе, короче, автополярно, т. е. существует полярное преобразование, которое сопоставляет каждой точке конического сечения его касательную в этой точке, и обратно.*

Доказательство. Рассмотрим в связке прямых круговой конус K_0, у которого прямой угол раствора — угол между противоположными образующими, т. е. лежащими в одной плоскости с осью конуса. Плоскость, проходящая через вершину перпендикулярно образующей, касается конуса по противоположной образующей. Стало быть, соответствие прямых и плоскостей по взаимной перпендикулярности сопоставляет образующим конуса его касательные плоскости; очевидно и обратное: касательным плоскостям сопоставляются образующие.

Тут образующей сопоставляется плоскость, касающаяся конуса вдоль другой образующей — противоположной. Но если заменим в связке все прямые на симметричные им относительно оси конуса, то образующие конуса заменятся на противоположные. В результате получим соответствие прямых и плоскостей, при котором каждой образующей конуса соответствует плоскость, касательная к конусу вдоль той же образующей.

Далее можно аффинным преобразованием всей связки превратить конус K_0 в любой данный конус второго порядка K.[9] В результате получим полярное соответствие, при котором каждой образующей ко-

[9]Конус K_0, у которого угол раствора прямой, задается в подходящих координатах уравнением $x^2 + y^2 - z^2 = 0$. Уравнение $x^2/a^2 + y^2/b^2 - z^2/c^2 = 0$ задает конус второго порядка общего вида. Преобразованием $x' = x/a$, $y' = y/b$, $z' = z/c$ он переводится в K_0.

нуса K соответствует его касательная плоскость, касающаяся вдоль этой образующей.

Рассматривая связку как проективную плоскость, мы получаем на ней полярное соответствие точек и прямых, при котором конус K представляет коническое сечение, у которого полученная полярность сопоставляет каждой точке касательную прямую в этой же точке.

Так как конус K — любой конус второго порядка, то он представляет любое коническое сечение.

Таким образом автополярность любого конического сечения доказана. $\square$

Полярность, связанная с коническим сечением. Выясним, как при данном коническом сечении K установить на всей плоскости то полярное соответствие, при котором точкам данного конического сечения K соответствуют касательные в них же.

Пусть K — данное коническое сечение. Его касательные в точках A, B пересекаются в некоторой точке P. Так как точки A, B соответствуют прямым a, b, то точка P, где пересекаются эти прямые, соответствует прямой p, проходящей через A и B. Точка P — это *полюс* прямой p, а прямая p — *поляра* точки P (рис. 78). Для наглядности можно представлять себе, что коническое сечение K — окружность. Так как

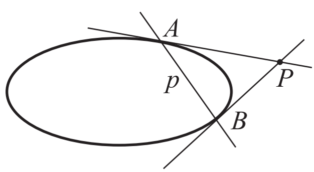


Рис. 78

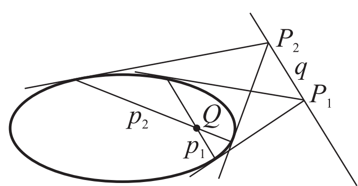


Рис. 79

через каждую точку P вне конического сечения — вне его внутренней области (как вне круга) — проходит две касательные к нему, то данное построение сопоставляет каждой внешней точке P прямую p, пересекающую коническое сечение.

Если P_1, P_2 — две точки вне конического сечения и q — проходящая через них прямая, то ей соответствует точка Q, где пересекаются поляры p_1, p_2 точек P_1, P_2. Эта точка лежит внутри конического сечения. Прямая q будет ее полярой.

Если точка P движется по прямой q, то ее поляра вращается вокруг точки Q (рис. 79).

В итоге каждой точке — вне или внутри конического сечения — сопоставляется поляра, и каждой прямой — полюс.

Полярность относительно окружности. Можно считать какую-либо прямую, не пересекающую коническое сечение, бесконечно удаленной прямой, а само коническое сечение — окружностью. Тогда установленное полярное соответствие связано с окружностью и выражает факты элементарной геометрии. Прежде всего можно заметить, что касательные к окружности, пересекающиеся в бесконечно удаленной точке, т. е. параллельные, касаются окружности в диаметрально противоположных точках. Прямая, проходящая через них, проходит через центр — это диаметральная прямая, или, как принято коротко говорить в теории кривых второго порядка, — диаметр. Таким образом, полярой p бесконечно удаленной точки P служит диаметр.

Когда бесконечно удаленная точка P движется по бесконечно удаленной прямой, т. е. когда направление параллельных прямых поворачивается, диаметр p вращается вокруг центра. Поэтому центр окружности является полюсом бесконечно удаленной прямой.

Теперь обратимся к обычным точкам вне круга. Полярное соответствие означает следующее.

Теорема 2. *Каждой точке P вне круга сопоставим прямую p, проходящую через те точки, где касательные, проведенные через P, касаются окружности круга. Тогда если точка P зачерчивает какую-нибудь прямую q, то соответствующая прямая p вращается вокруг одной точки Q* (совершая полный оборот; рис. 80). Тем самым выполняется и обратное: *если какая-либо прямая p вращается вокруг некоторой заранее выбранной точки Q, то соответствующая этой прямой точка P зачерчивает прямую q, соответствующую точке Q.* $\square$

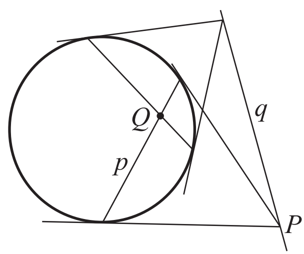


Рис. 80

Эту теорему можно считать прежде всего теоремой о касательных к окружности, так как она говорит о прямых, проходящих через точки касания. Стоит внимательно рассмотреть данную теорему и попытаться доказать ее «в лоб» — средствами элементарной геометрии, вовсе не ссылаясь на соображения двойственности из проективной геометрии.

§ 4. Проективное пространство

Проективное пространство можно получить, пополняя евклидово пространство бесконечно удаленными точками, которые образуют бесконечно удаленную плоскость: каждая другая плоскость пересекает ее по своей бесконечно удаленной прямой. Параллельные плоскости пересекаются по бесконечно удаленной прямой — она у них общая, так же как параллельные прямые пересекаются в их общей бесконечно удаленной точке.

Пространство, пополненное таким образом бесконечно удаленной плоскостью, становится проективным пространством, когда бесконечно удаленные его элементы — точки, прямые, плоскость — полагаются «равноправными» с соответствующими обычными его элементами. Точно это можно определить с помощью проективных преобразований.

Проективным преобразованием пополненного пространства называется такое его взаимно однозначное отображение на себя, при котором прямые отображаются на прямые, плоскости — на плоскости, без различия между обычными и бесконечно удаленными.

Благодаря таким преобразованиям пополненное пространство и становится проективным, где уже нет особых бесконечно удаленных элементов. Поэтому можно дать такое его определение.

Проективным пространством называется пространство, пополненное бесконечно удаленными элементами, если в нем введены проективные преобразования.

Свойство фигур называется *проективным*, если оно сохраняется при любых проективных преобразованиях — инвариантно относительно этих преобразований.

Проективной геометрией в пространстве называется теория — область геометрии, — в которой изучаются проективные свойства фигур.

Подобно тому как проективную плоскость можно представить связкой прямых, так и проективное пространство можно представить связкой прямых, но уже в **четырехмерном** пространстве. Прямые связки служат при этом точками проективного пространства.

Проективным преобразованием является отображение связки на себя — ее прямых на прямые, — вызываемое любым аффинным преобразованием объемлющего евклидова пространства, сохраняющим центр связки.

Прямые проективного пространства представляются (двумерными) плоскостями связки, т. е. проходящими через ее центр, а плоскость

представляется **трехмерными плоскостями** связки. В каждой из них содержится трехмерная связка, которая и представляет ее — эту проективную плоскость.

Совершенно так же, как при изображении проективной плоскости связкой прямых, вводятся однородные координаты; теперь их четыре: x_1, x_2, x_3, x_4. Они определены с точностью для любого отличного от нуля общего множителя. Как говорят, определено их отношение $(x_1 : x_2 : x_3 : x_4)$. Оно задает прямую связки и тем самым — точку проективного пространства. (В объемлющем пространстве, содержащем связку, x_1, x_2, x_3, x_4 — это какие-либо аффинные координаты с началом в центре связки. На прямой связки они отличаются только общим множителем.)

Аффинное преобразование объемлющего пространства, сохраняющее центр связки, представляется в координатах линейным преобразованием (с определителем, отличным от нуля). Это линейное преобразование, в котором координаты x_1, x_2, x_3, x_4 рассматриваются как однородные координаты точек, и представляет проективное преобразование проективного пространства.

Как от проективной плоскости можно перейти обратно к аффинной, фиксируя какую-нибудь прямую в качестве бесконечно удаленной, так и от проективного пространства можно перейти к аффинному, фиксируя какую-нибудь плоскость и приняв ее за бесконечно удаленную. Аффинные преобразования пространства — это его проективные преобразования, при которых некоторая плоскость отображается сама на себя в качестве бесконечно удаленной.

Так же как в случае проективной плоскости, это дает средство доказательства теорем проективной геометрии о фигурах, не пересекающих все плоскости пространства.

Пусть фигура F не имеет с плоскостью α общих точек. Примем эту плоскость за бесконечно удаленную; фигура F окажется в аффинном пространстве, и ее подлежащее доказательству свойство окажется ее аффинным свойством. Можно даже принять пространство за евклидово, введя в нем соответствующее расстояние. Свойство фигуры F можно будет выразить и доказать в этих условиях — в евклидовом пространстве. Но так как это свойство по его формулировке проективное, то оно и будет доказано в этом его качестве. Проективное — значит, не изменяется при любых проективных преобразованиях; значит, доказав его в дополнительных условиях, мы доказали его во всей общности.

В проективном пространстве можно рассмотреть поверхность второго порядка, так же как на проективной плоскости рассмотрены

КВП. Проведите сами проективную классификацию ПВП. Обратите внимание на то, что эллипсоид, двуполостный гиперболоид и эллиптический параболоид попадают в один класс — представляют одинаковые поверхности, только по-разному расположены относительно плоскости, принятой за бесконечно удаленную. Другой гиперболоид — однополостный — и гиперболический параболоид тоже попадают в один класс. Очевидно, наличие прямолинейных образующих есть проективное свойство. Цилиндры — это конусы с бесконечно удаленной вершиной...

Разберите все случаи в зависимости от расположения бесконечно удаленной плоскости.

Теорема Дезарга. Докажем теорему Дезарга в проективном пространстве. Отличие от обычной стереометрии только в том, что две прямые, лежащие в одной плоскости, всегда пересекаются.

Пусть у двух треугольников вершины поставлены в соответствие, и тем самым поставлены в соответствие и стороны. Теорема утверждает: если прямые, соединяющие соответственные вершины, пересекаются в одной точке, то точки пересечения прямых, содержащих соответственные стороны, лежат на одной прямой, и обратно: если точки пересечений прямых, содержащих соответственные стороны, лежат на одной прямой, то прямые, проходящие через соответственные вершины, пересекаются в одной точке.

Собственно говоря, треугольники здесь ни при чем, а речь идет о двух тройках точек и проходящих через них прямых; так и сформулируем теорему.

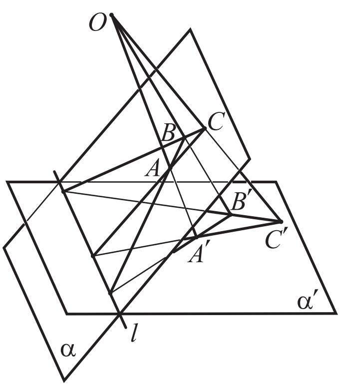


Рис. 81

Теорема (Дезарг). *Пусть точки A, B, C не лежат на одной прямой, как и точки A', B', C'. Тогда если прямые AA', BB', CC' пересекаются в одной точке, то точки пересечения прямых AB с $A'B'$, AC с $A'C'$, BC с $B'C'$ лежат на одной прямой. И обратно: если точки пересечения прямых AB с $A'B'$, AC с $A'C'$, BC с $B'C'$ лежат на прямой, то прямые AA', BB', CC' пересекаются в одной точке.*

Доказательство при условии, что точки A, B, C и A', B', C' лежат в разных плоскостях α, α'.

Докажем первое утверждение теоремы. Пусть прямые AA', BB' пересекаются в точке O. Тогда они лежат в одной плоскости и в той же плоскости лежат прямые AB, $A'B'$. Следовательно, они пересекаются. Но эти прямые лежат в плоскостях α, α', а поэтому их точка пересечения лежит на прямой, по которой эти плоскости пересекаются (рис. 81).

Тот же вывод верен для прямых AC, $A'C'$ и BC, $B'C'$. Первая часть теоремы доказана.

Пусть прямые AB и $A'B'$ пересекаются. Следовательно, они лежат в одной плоскости, а потому прямые AA', BB' лежат в той же плоскости и, стало быть, пересекаются в некоторой точке O.

Совершенно так же убедимся, что прямые AA', CC' тоже пересекаются в одной точке O_1, а прямые BB', CC' — в какой-то точке O_2. Тогда прямая AA' проходит через точки O, O_1, BB' — через O, O_2, CC' — через O_1, O_2. Поэтому если бы точки O, O_1, O_2 не лежали на одной прямой, то они определяли бы плоскость и вместе с ними все точки A, $\ldots$, C' лежали в одной плоскости вопреки условию. Следовательно, точки O, O_1, O_2 лежат на одной прямой.

Но O, O_1 лежат на прямой AA', а потому O_2 тоже лежит на прямой AA'. А так как в O_1 пересекаются прямые BB', CC', то выходит, что все три прямые AA', BB', CC' пересекаются в одной точке, что и требовалось доказать.

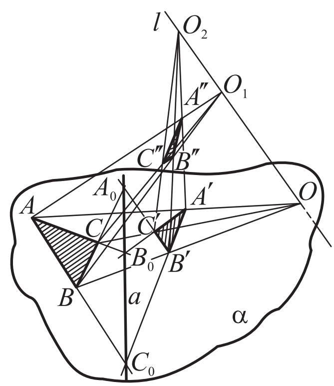


Рис. 82

Случай, когда точки A, B, C, A', B', C' лежат в одной плоскости, был рассмотрен раньше путем сведения к элементарной геометрии посредством введения бесконечно удаленной прямой. Но можно обойтись без этого с помощью вспомогательного построения. Докажем таким приемом первую часть теоремы Дезарга в случае, когда точки A, B, C, A', B', C' лежат в одной плоскости.

Итак, пусть эти точки лежат в плоскости α, и пусть прямые AA' BB', CC' пересекаются в точке O (рис. 82).

Проведем через нее прямую l, пересекающую плоскость α, и возьмем на ней две точки O_1, O_2. Прямые O_1A, O_2A' лежат в плоскости, проходящей через прямые l и AA'. Поэтому они пересекаются в некоторой точке A''. Совершенно так же убедимся, что прямые

O_1B, O_2B' и O_1C, O_2C' пересекаются в некоторых точках B'', C''.

Так мы получаем две тройки точек A, B, C и A'', B'', C'', а также A', B', C' и A'', B'', C'', лежащие в разных плоскостях, для которых выполнено условие теоремы (прямые AA'', BB'' и $A'A''$, $B'B''$ пересекаются в точках O_1 и O_2).

Плоскость α пересекается с плоскостью точек A'', B'', C'' по некоторой прямой a. Эту прямую прямая $A''B''$ пересекает в некоторой точке C_0. Но по доказанному прямые AB, $A'B'$ пересекаются с $A''B''$ на прямой a. Тем самым они пересекаются друг с другом в той же точке C_0. Аналогично, на прямой a лежат точка A_0 пересечения прямых BC и $B'C'$ и точка B_0 пересечения прямых AC и $A'C'$.

Таким образом, все три пары прямых AB, $A'B'$ и т. д. пересекаются в точках, лежащих на одной прямой, что и требовалось доказать.

Доказательство второй части теоремы Дезарга для того же случая мы оставим читателю. □

Глава V
МНОГОМЕРНАЯ ЕВКЛИДОВА ГЕОМЕТРИЯ

§ 1. Аксиомы n-мерного пространства. Векторы и координаты

Идея многомерного евклидова пространства очень проста. На прямой точка задается одной координатой; на плоскости — двумя координатами, в обычном (трехмерном) пространстве — тремя координатами. Дальше можно мыслить четырехмерное пространство, где точка задается четырьмя координатами, и вообще n-мерное пространство, где точка задается n координатами, а натуральное число n может быть любым данным.

Координаты можно представить себе как прямоугольные, так что n координатам соответствует n взаимно перпендикулярных прямых — осей координат. В четырехмерном пространстве их четыре — к осям x, y, z добавляется ось t; оси пересекаются в начале координат O. Через каждые две оси проходит плоскость, так что есть, в частности, две плоскости xy и zt. У них только одна общая точка — начало координат. Таким образом, та аксиома, что две плоскости с общей точкой пересекаются по прямой, отпадает. Ее нужно заменить другой. А в остальном, можно сказать, аксиомы n-мерной геометрии совершенно сходны с аксиомами стереометрии. Сформулируем их.

Аксиоматика n-мерной евклидовой геометрии.

Основные объекты: 1) *точки*, 2) *отрезки*, 3) *плоскости.*

Основные отношения: 1) *точка служит концом отрезка,* 2) *точка лежит на отрезке,* 3) *точка принадлежит плоскости,* 4) *отрезки равны.*

Аксиомы делятся на: 1) *линейные,* 2) *плоскостные,* 3) *пространственные.* Линейные аксиомы те же, что в планиметрии; плоскостные аксиомы те же, но, как и в стереометрии, их нужно относить к каждой плоскости. Пространственных аксиом три:

Пр. 1 (аксиома плоскости). *Каждые три точки принадлежат плоскости.*

Пр. 2 (аксиома общей прямой). *Если у двух плоскостей есть две общие точки, то они пересекаются по прямой.*

Пр. 3 (аксиома числа измерений). *Существует n и не более взаимно перпендикулярных прямых, т. е. пересекающихся в общей точке и образующих попарно прямые углы* (определение прямого угла то же, что в планиметрии).

Можно дать определение: n-мерная евклидова геометрия — это теория, которая строится на указанной аксиоматике; n-мерное евклидово пространство — это множество некоторых элементов — «точек», в котором выполнены указанные аксиомы, так что в нем определены отрезки и отношение их к точкам и друг к другу[10].

Если $n = 2$, то пространство — это плоскость, а теория — это планиметрия; аксиомы Пр. 1, 2 формально выполнены, но оказываются бессодержательными.

При $n > 2$ существуют точки, не лежащие в одной плоскости (как высказывалось в аксиоме стереометрии), потому что на плоскости нет трех взаимно перпендикулярных прямых.

Совершенно так же, как в стереометрии, в общем случае $n \geqslant 3$ доказываются теоремы.

1. *Через каждые две точки проходит прямая, и притом только одна* (впрочем, это следует из одних линейных аксиом).

2. *Если прямая имеет с плоскостью две общие точки, то она вся содержится в этой плоскости. Точно так же отрезок, соединяющий две точки плоскости, содержится в ней.*

---

[10] Можно сказать: в этом множестве — в пространстве — выделены (определены) подмножества, называемые «отрезками», причем у каждого отрезка из принадлежащих ему точек отмечены две — «концы», и между отрезками установлено отношение «равенства», все это подчиняется сформулированным аксиомам.

Отсюда и из плоскостных аксиом, так же как в § 6 гл. I ч. 2, вытекает важное заключение.

На всякой плоскости выполняется планиметрия.

Далее, так же как в § 6 гл. I ч. 2, доказывается: через три точки, не лежащие на одной прямой, так же как и через две пересекающиеся прямые, как и через прямую и не лежащую на ней точку, проходит плоскость, и притом только одна.

(Заметим, что при $n = 3$ из высказанных аксиом следует, что две плоскости, имеющие хотя бы одну общую точку, пересекаются по прямой, как это сказано в аксиомах стереометрии.)

Векторы и координаты. Опираясь на сформулированные начальные выводы, можно определить векторы и операции с ними буквально так же, как это сделано в трехмерном пространстве. Прослеживая определения и выводы главы о векторах (ч. 1, гл. III), можно убедиться, что в них число измерений пространства не играет роли нигде, кроме как в разложении векторов на составляющие. (Читателю полезно пройти по главе о векторах и убедиться в сказанном.) Таким образом, мы можем в n-мерном пространстве применить без всяких изменений понятие вектора, сложение векторов, умножение на число и скалярное произведение.

Пользуясь этим, введем в n-мерном пространстве прямоугольные координаты.

По аксиоме числа измерений в пространстве есть n и не более взаимно перпендикулярных прямых e_1, e_2, …, e_n. На них зададим единичные векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, …, $\boldsymbol{e}_n$. Они взаимно ортогональны, поэтому их скалярные произведения равны нулю, и мы имеем равенство

$$\boldsymbol{e}_i \boldsymbol{e}_j = \begin{cases} 1, & \text{если} \quad i = j, \\ 0, & \text{если} \quad i \neq j. \end{cases} \tag{1}$$

Покажем, что *всякий вектор* $\boldsymbol{a}$ *представляется через эти векторы:*

$$\boldsymbol{a} = \sum_{i=1}^{n} \boldsymbol{e}_i (\boldsymbol{a}\boldsymbol{e}_i), \tag{2}$$

где коэффициенты $(\boldsymbol{a}\boldsymbol{e}_i)$ — координаты вектора $\boldsymbol{a}$ относительно векторов $\boldsymbol{e}_1$, …, $\boldsymbol{e}_n$ — представляют собой скалярные произведения $\boldsymbol{a}$ на векторы $\boldsymbol{e}_i$.

Для доказательства возьмем какой угодно вектор $\boldsymbol{a}$ и рассмотрим

вектор
$$\boldsymbol{c} = \boldsymbol{a} - \sum_{i=1}^{n} \boldsymbol{e}_i(\boldsymbol{a}\boldsymbol{e}_i). \tag{3}$$
Найдем его скалярное произведение на какой-либо из векторов $\boldsymbol{e}_k$:
$$\boldsymbol{c}\boldsymbol{e}_k = \boldsymbol{a}\boldsymbol{e}_k - \sum_{i=1}^{n}(\boldsymbol{e}_k\boldsymbol{e}_i)(\boldsymbol{a}\boldsymbol{e}_i). \tag{4}$$
Но ввиду равенств (1) все произведения $\boldsymbol{e}_k\boldsymbol{e}_i$ равны нулю, кроме $\boldsymbol{e}_k\boldsymbol{e}_k = 1$. Поэтому в правой части равенства (4) остается только член $(\boldsymbol{e}_k\boldsymbol{e}_k)(\boldsymbol{a}\boldsymbol{e}_k) = \boldsymbol{a}\boldsymbol{e}_k$, и мы получаем
$$\boldsymbol{c}\boldsymbol{e}_k = \boldsymbol{a}\boldsymbol{e}_k - \boldsymbol{a}\boldsymbol{e}_k = 0.$$

Таким образом, выходит, что либо вектор $\boldsymbol{c}$ равен нулю, либо он ортогонален всем $\boldsymbol{e}_k$. Но последнее исключено.

Действительно, по аксиоме Пр. 3 есть не более n взаимно перпендикулярных прямых $e_1, \ldots, e_n$, пересекающихся в общей точке O. Если бы вектор $\boldsymbol{c}$ был отличен от нуля, то отложив его от точки O и проведя вдоль него прямую, мы получили бы еще $(n+1)$-ю прямую, перпендикулярную всем прямым $e_1, e_2, \ldots, e_n$. Это противоречит аксиоме. Следовательно, $\boldsymbol{c} = \boldsymbol{0}$, и тем самым выполняется равенство (2): всякий вектор $\boldsymbol{a}$ представи́м в таком виде через векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. $\square$

Пусть теперь M — произвольная точка и $\overrightarrow{OM} = \boldsymbol{r}$ — ее радиус-вектор. Мы определяем ее координаты x_i как скалярные произведения
$$x_i = \boldsymbol{r}\boldsymbol{e}_i = \overrightarrow{OM} \cdot \boldsymbol{e}_i \quad (i = 1,\ 2,\ \ldots,\ n). \tag{5}$$
Это — координаты относительно основного n-векторника $O\boldsymbol{e}_1\boldsymbol{e}_2 \ \ldots \ \boldsymbol{e}_n$; осями координат служат взаимно перпендикулярные прямые, существующие согласно аксиоме Пр. 3.

Так в n-мерном пространстве оказываются введенными прямоугольные координаты. Каждой точке M однозначно соответствует набор ее координат $(x_1, x_2, \ldots, x_n)$, и разным точкам — разные наборы, так что точки однозначно задаются своими координатами.

Действительно, допустим, точки M, N имеют одни и те же координаты $x_1, x_2, \ldots, x_n$. Это значит, согласно определению координат формулой (5), что $\overrightarrow{OM} \cdot \boldsymbol{e}_i = \overrightarrow{ON} \cdot \boldsymbol{e}_i$ $(i = 1,\ 2,\ \ldots,\ n)$, т. е.
$$(\overrightarrow{OM} - \overrightarrow{ON})\boldsymbol{e}_i = 0.$$

Так что либо вектор $\overrightarrow{OM}-\overrightarrow{ON}$ нулевой, либо он перпендикулярен всем $\boldsymbol{e}_i$. Последнее, однако, противоречило бы аксиоме числа измерений. Следовательно, $\overrightarrow{OM} - \overrightarrow{ON} = \mathbf{0}$, т. е. точки M и N совпадают — это одна точка, что и требовалось доказать.

Убедимся еще, что каждому набору чисел $x_1, x_2, \ldots, x_n$ соответствует точка с такими координатами. Она определяется как конец радиус-вектора

$$\overrightarrow{OM} = \sum \boldsymbol{e}_i x_i. \tag{6}$$

Умножая скалярно на любой вектор $\boldsymbol{e}_k$, получаем ввиду формул (1)

$$\overrightarrow{OM}\,\boldsymbol{e}_k = \sum (\boldsymbol{e}_k \boldsymbol{e}_i) x_i = x_k,$$

т. е. x_k и есть координата точки M.

Подведем итог, формулируя его как доказанную теорему.

Теорема 1 (о координатах). *В n-мерном пространстве можно ввести координаты, в которых осями служат те n взаимно перпендикулярных прямых, существование которых утверждает аксиома числа измерений. Если $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — единичные векторы вдоль этих прямых и O — точка их пересечения, то координаты точки M определяются из скалярного произведения:*

$$x_i = \boldsymbol{e}_i \overrightarrow{OM} \quad (i = 1, \ldots, n).$$

Этим между точками и наборами чисел $x_1, \ldots, x_n$ устанавливается взаимно однозначное соответствие. $\square$

Эти же координаты получаются, как и в трехмерном пространстве, если проектировать точку M на оси. Именно, проводим через данную точку M и k-ю ось — прямую e_k — плоскость и в этой плоскости спроектируем точку M на эту прямую. Координата проекции и будет равна x_k.

Теперь обратимся к длине. Длина $|\boldsymbol{a}|$ вектора $\boldsymbol{a}$ связана с его скалярным произведением самого на себя:

$$|\boldsymbol{a}|^2 = \boldsymbol{a}^2.$$

Вместе с тем по формуле (2)

$$\boldsymbol{a} = \sum \boldsymbol{e}_i a_i, \quad a_i = (\boldsymbol{a}\boldsymbol{e}_i).$$

Поэтому, выполняя скалярное возведение в квадрат и пользуясь соотношениями (1), получим (все произведения $\boldsymbol{e}_i \boldsymbol{e}_j$, $i \neq j$, исчезают и остаются только квадраты $\boldsymbol{e}_i^2 = 1$)

$$|\boldsymbol{a}|^2 = \boldsymbol{a}^2 = \left(\sum \boldsymbol{e}_i a_i\right)^2 = \sum a_i^2. \tag{7}$$

Отсюда выводится:

Теорема 2 (о длине отрезка). *Длина отрезка с концами* $A(x_1, \ldots, x_n)$, $B(x_1', \ldots, x_n')$, *или, что то же, расстояние* $|AB|$ *между этими точками выражается формулой*

$$|AB| = \sqrt{(x_1 - x_1')^2 + \ldots + (x_n - x_n')^2}. \tag{8}$$

Доказательство. Если O, как и выше, начало координат, то по формуле (6)

$$\overrightarrow{OA} = \sum_{i=1}^{n} \boldsymbol{e}_i x_i, \quad \overrightarrow{OB} = \sum_{i=1}^{n} \boldsymbol{e}_i x_i'.$$

Поэтому

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \sum_{i=1}^{n} \boldsymbol{e}_i (x_i' - x_i).$$

Откуда по формуле (7) получаем

$$|AB|^2 = \sum_{i=1}^{n} (x_i' - x_i)^2,$$

что и дает формулу (8). $\square$

Шар и сфера определяются в n-мерном пространстве буквально так же, как в трехмерном. Повторите эти определения; напишите уравнение сферы и неравенство, задающее шар, в n-мерном пространстве.

§ 2. Прямые и плоскости разного числа измерений

Параллельные прямые определяются в многомерном пространстве так же, как в трехмерном, и так же доказываются теоремы:

1. *Через точку, не лежащую на данной прямой, проходит прямая, ей параллельная, и притом только одна.*

2. *Две прямые, параллельные одной и той же прямой, параллельны друг другу.*

Доказательство такое же, как в § 3 гл. II ч. 2.

Доказательства других теорем из стереометрии, данные в том же параграфе, также переносятся на случай пространства любого числа измерений, поскольку они основаны на применении векторов. Убедитесь в этом, проследив эти доказательства.

Однако есть разница в понятии параллельных плоскостей. В трехмерном пространстве, в отличие от плоскости, прямые не считаются параллельными, если они всего лишь не пересекаются: они могут скрещиваться. Так же в многомерном пространстве параллельность плоскостей или прямой и плоскости не определяется тем, что они только не пересекаются.

Для определения параллельности воспользуемся понятием о параллельном переносе. Он определяется буквально так же, как на плоскости или в трехмерном пространстве. Это отображение, при котором точки X переходят в такие точки Y, что все векторы $\overrightarrow{XY}$ равны.

Иными словами, все точки переносимой фигуры переносятся на один и тот же вектор. Все свойства переноса, установленные в § 2 главы о наложениях, равно относятся к многомерному пространству.

Теперь даем определение. Прямые или плоскости *параллельны*, если одна переводится в другую переносом. Прямая *параллельна* плоскости, если она не лежит в этой плоскости, но ее можно отобразить в эту плоскость переносом.

Легко убедиться, что эти определения в случае трехмерного пространства равносильны обычным. (Докажите это.) Но, скажем, в четырехмерном пространстве E^4 прямая может «скрещиваться» с плоскостью — не пересекать ее и не быть ей параллельной. Например, возьмем в E^4 две плоскости P, Q, имеющие только одну общую точку O,[11] и в одной из них — прямую a, не проходящую через O. Она не пересекает другую плоскость и не помещается в нее переносом (убедитесь в этом, рассмотрев проходящую через O прямую $b||a$).

[11] Например, плоскость P, проходящую через оси e_1, e_2, а Q — через оси e_3, e_4.

Точно так же две плоскости, имеющие общий перпендикуляр, могут не быть параллельными, как в трехмерном пространстве две прямые с общим перпендикуляром не обязаны быть параллельными.

Через данную точку плоскости в более чем трехмерном пространстве проходит не одна перпендикулярная ей прямая. В четырехмерном пространстве перпендикуляры к плоскости образуют плоскость. Так, перпендикуляры в точке O к плоскости, проходящей через оси координат e_1, e_2, заполняют плоскость, проходящую через e_3, e_4. (Убедитесь в этом.)

Общее понятие плоскости. Если пространство более чем трехмерное (не выполняется аксиома пересечения плоскостей), то в нем наряду с обычными двумерными плоскостями есть другие фигуры, также называемые плоскостями.

Определение. *Плоскостью*, вообще, называется такая фигура, которая содержит по крайней мере три точки, не лежащие на одной прямой, и вместе с каждыми двумя своими точками содержит всю содержащую их прямую, но при этом не является всем пространством. Обычная двумерная плоскость, очевидно, подпадает под это определение[12].

Теорема 1. *Всякая недвумерная плоскость представляет собою евклидово пространство, т. е. в ней выполняются аксиомы, определяющие евклидово пространство, с теми же двумерными плоскостями, какие имеются в объемлющем пространстве*[13].

Через всякие три точки любой данной плоскости проходит содержащаяся в ней двумерная плоскость.

Доказательство. Пусть P — данная плоскость в каком-либо евклидовом пространстве и A, B, C — какие-либо три ее точки. По определению плоскости она содержит прямую AB, но не сводится к ней. Поэтому если точка C оказалась на этой прямой, то можно взять точку D, не лежащую на прямой AB. Если C не лежит на AB, то D и есть C.

Плоскость P, по ее определению, содержит прямые AB и AD. Через

[12] Обобщая, можно определить плоскость общим условием, что она вместе с любыми двумя точками содержит и проходящую через них прямую. Тогда «плоскостью» может быть: 1) пустое множество, 2) одна точка, 3) прямая, 4) все пространство. Первые три возможности исключаются требованием, что плоскость содержит по крайней мере три точки не на прямой; четвертая — исключена последней оговоркой в определении.

[13] Вообще, фигура может быть евклидовым двумерным пространством и в этом смысле плоскостью, вовсе не будучи плоскостью в объемлющем пространстве, как, например, параболический цилиндр с точки зрения его внутренней геометрии.

эти прямые проходит двумерная плоскость Q, образованная пересекающими их прямыми. Все эти прямые будут содержаться в плоскости P (в силу определения плоскости). Следовательно, в P содержится двумерная плоскость Q, проходящая через данные точки A, B, C, что и требовалось доказать.

Итак, первая аксиома евклидова пространства выполнена. Вторая выполняется очевидным образом, поскольку она выполняется в объемлющем пространстве для любых двух плоскостей с двумя общими точками. $\square$

Число измерений (или, что то же, *размерность*) плоскости можно определить, как в пространстве, числом взаимно перпендикулярных прямых. Прямую можно считать одномерной плоскостью.

Плоскость P можно назвать *параллельной* плоскости Q, если существует перенос, помещающий ее в плоскость Q, т. е. такой перенос t, что $tP \subset Q$. Мы пишем: $P||Q$. (Заметим, что плоскости P и Q могут быть разных размерностей, поэтому одновременно $Q||P$, только если $tP = Q$, так что Q перемещается в P обратным переносом t^{-1}).

В четырехмерном пространстве есть трехмерные плоскости: каждая из них — это трехмерное евклидово пространство. Если две такие плоскости имеют общую точку, то их пересечение представляет двумерную плоскость. (Рассмотрите на примере плоскостей, «натянутых» на оси координат.)

Часть 4
ДИФФЕРЕНЦИАЛЬНАЯ ГЕОМЕТРИЯ

В четвертой части мы изучаем кривые и поверхности в трехмерном пространстве. Основной инструмент для исследования кривых — это естественная параметризация. С ее помощью мы даем первоначальные определения большинства вводимых дифференциально-геометрических понятий, а затем уже приводим чисто геометрические определения. Глава о кривых заканчивается «натуральными уравнениями», описывающими вид кривой вне зависимости от ее расположения в пространстве.

Для исследования поверхностей мы рассматриваем лежащие на них кривые: с их длиной и кривизной естественно связываются первая и вторая основные формы поверхности. В конце второй главы рассматриваются локально кратчайшие кривые — геодезические. Это аналоги прямых на плоскости, — они играют важную роль при изучении внутренней геометрии на искривленной поверхности.

Глава I
ДИФФЕРЕНЦИАЛЬНАЯ ГЕОМЕТРИЯ КРИВЫХ

§ 1. Элементарные кривые на плоскости и в пространстве. Способы их задания

Для изучения кривых необходимо иметь их точное определение. Мы ограничимся здесь тем, что дадим определение элементарной кривой. Множество C (на плоскости или в пространстве) называется *элементарной кривой*, если оно является образом отрезка при некотором непрерывном взаимно однозначном (т. е. инъективном) отображении этого отрезка в плоскость или в пространство.

Примеры. Простейшие элементарные кривые — это отрезки прямых, дуги окружностей, эллипсов, графики непрерывных функций,

заданных на отрезке, и т. п. Простые конечнозвенные ломаные тоже будут элементарными кривыми в нашем определении (рис. 1).

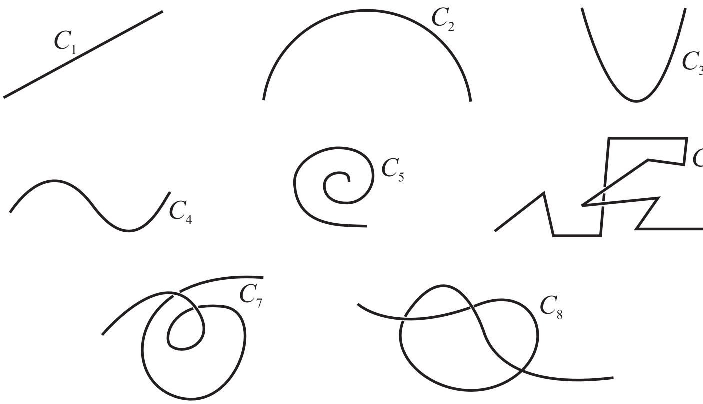


Рис. 1

Образы концов отрезка называются *концами* элементарной кривой, а образ любого отрезка, содержащегося в исходном отрезке, называется *дугой*. Очевидно, что всякая дуга элементарной кривой сама является элементарной кривой.

Если элементарная кривая C есть образ отрезка $[a,\ b]$ при взаимно однозначном и непрерывном отображении $F\colon [a,\ b] \to \mathbf{R}^3$, то положение любой точки P на кривой C определяется одним-единственным числом $t \in [a,\ b]$, образом которого эта точка является: $P = F(t)$. Переменная t называется *параметром* кривой C. Разным значениям параметра соответствуют различные точки кривой C. Отображение F будем называть *параметризацией* кривой C. У одной и той же элементарной кривой может быть много различных параметризаций. Кривую, снабженную параметризацией, будем называть *параметризованной* кривой (рис. 2).

Фиксируем систему координат. Пусть точка $P = F(t)$ имеет координаты x, y, z. При изменении параметра t они тоже будут меняться — каждая координата является некоторой функцией от t:

$$\begin{cases} x = f_1(t), \\ y = f_2(t), \\ z = f_3(t), \end{cases} \tag{1}$$

где $f_1,\ f_2,\ f_3$ — непрерывные числовые функции, заданные на отрезке $[a,\ b]$ (рис. 3).

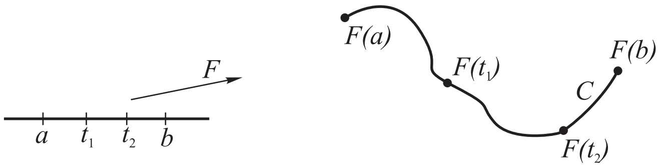


Рис. 2

Функции $f_1,\ f_2,\ f_3$ полностью описывают параметризацию F и называются ее *координатными функциями*. Соотношения $x = f_1$, $y = f_2$, $z = f_3$ называются *уравнениями* параметризованной кривой C.

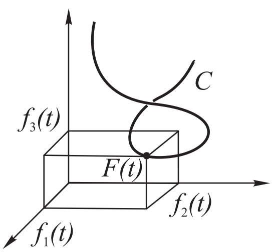


Рис. 3

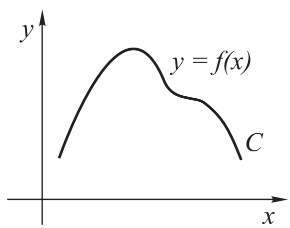


Рис. 4

Если $f\colon [a,\ b] \to \boldsymbol{R}$ — непрерывная функция, то ее график является плоской элементарной кривой, допускающей параметризацию $x = t$, $y = f(t)$. Как множество кривая C задается уравнением $y = f(x)$. Такое задание кривой называется *явным* (рис. 4). Пространственная кривая допускает *явное задание*, если она обладает параметризацией $x = t$, $y = f(t)$, $z = g(t)$. Такая кривая как множество может быть задана системой уравнений $y = f(x)$, $z = g(x)$.

Не все кривые допускают явное задание.

Пример. Любая дуга окружности, большая 180°, не допускает явного задания.

Для нас основной интерес будут представлять кривые, обладающие параметризацией с некоторыми дополнительными свойствами. Пусть $F\colon [a,\ b] \to \boldsymbol{R}^3$ — параметризация кривой C, а $f_1,\ f_2,\ f_3$ — ее координатные функции. Параметризация F называется *регулярной*, если, во-

первых, функции f_1 f_2, f_3 гладкие (т. е. достаточное число раз непрерывно дифференцируемые) и, во-вторых, при каждом значении параметра $t \in [a,\ b]$ производная по крайней мере одной из этих функций не обращается в нуль. (Последнее условие удобно записать в ви-

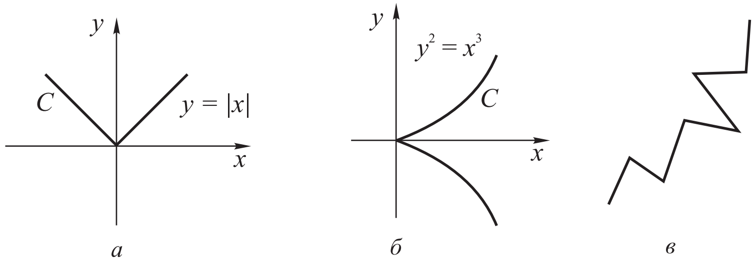


Рис. 5

де: $(f_1'(t))^2 + (f_2'(t))^2 + (f_3'(t))^2 \neq 0$.) Кривая, обладающая регулярной параметризацией, называется *гладкой* (ср. рис. 5). При необходимости мы будем без дополнительных оговорок требовать *повышенной гладкости*, т. е. существования у функций f_1, f_2, f_3 непрерывных производных до n-го порядка включительно при некотором $n \geqslant 2$.

Из теоремы о неявной функции следует, что в окрестности каждой своей точки гладкая кривая допускает явное задание. Другими словами, всякая достаточно малая дуга гладкой кривой является графиком некоторого гладкого отображения (в подходящей системе координат).

Пример. Рассмотрим отображение $[a,\ b] \xrightarrow{f} \boldsymbol{R}^2$, заданное формулой: $f(x) = (\cos x,\ \sin x)$. Его графиком в $\boldsymbol{R}^3$ будет отрезок *винтовой линии*. Он лежит на цилиндре радиуса 1 с осью Ox (рис. 6).

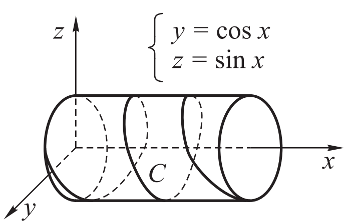


Рис. 6

Рассмотрим произвольное биективное непрерывное, а значит, и монотонное отображение φ некоторого отрезка $[c,\ d]$ в отрезок $[a,\ b]$: $\tau \longmapsto t = \varphi(\tau)$.

Если F: $[a,\ b] \to \boldsymbol{R}^3$ — параметризация кривой C, то сквозное отображение G: $[c,\ d] \to \boldsymbol{R}^3$, определяемое по формуле $G(\tau) = F(\varphi(\tau))$, также будет параметризацией кривой C. Говорят, что она получается из параметризации F при помощи *замены параметра* $t = \varphi(\tau)$. Если $\varphi(a) = c$ и $\varphi(b) = d$, то φ есть монотонно возрастающая функция, а

если $\varphi(a) = d$, $\varphi(b) = c$, то φ — монотонно убывающая функция (рис. 7).

Каждая параметризация определяет некоторый *порядок* точек на кривой. Если две параметризации связаны возрастающей заменой параметра, то они определяют один и тот же порядок, а если они связаны убывающей заменой параметра — то разный. Чтобы фиксировать порядок точек на кривой, достаточно указать *начальную* и *конечную* точки кривой (обе они являются концами кривой).

Элементарную кривую, у которой фиксированы начальная и конечная точки, назовем *ориентированной*. Всякая дуга ориентированной кривой сама ориентирована. В дальнейшем мы будем под *кривой* обычно подразумевать элементарную ориентированную кривую. Ориентацию удобно указывать стрелкой (рис. 8).

Замечание. Для того чтобы параметризация $G(\tau)$, полученная из регулярной параметризации $F(t)$ при помощи замены $t = \varphi(\tau)$, также была регулярной, необходимо и достаточно, чтобы замена была *неособой*, т. е. чтобы функция $\varphi(\tau)$ была непрерывно дифференцируемой, а ее производная $\varphi'(\tau)$ нигде не обращалась в нуль (она будет всюду положительной или всюду отрицательной). (Докажите это!)

До сих пор мы рассматривали кривую как множество точек или фигуру на плоскости или в пространстве. При параметризации кривой параметр играет роль координаты в

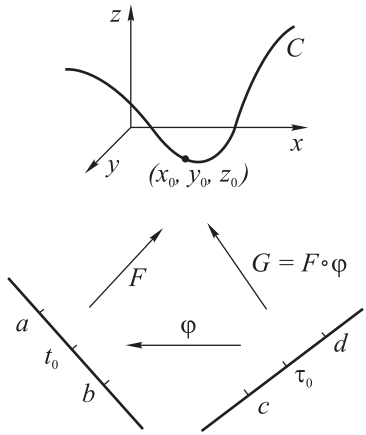


Рис. 7

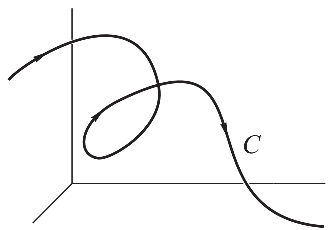


Рис. 8

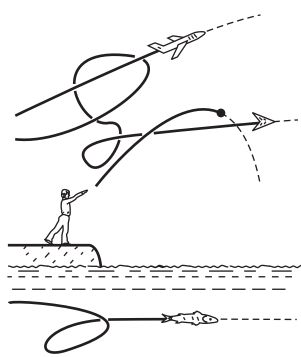


Рис. 9

этом множестве. (Координатной функцией служит функция, обратная параметризации.) Возможен и другой, очень плодотворный взгляд на кривую, как на траекторию движущейся материальной точки — птицы, рыбы, самолета и т. п. (рис. 9). (Например, летящий по инерции камень описывает параболу.) Здесь роль параметра играет время, прошедшее, например, от начала движения. Этот подход позволяет использовать в геометрии такие понятия из механики, как скорость, ускорение, путь и т. п.

§ 2. Вектор-функции одного переменного

Рассмотренные нами в § 1 способы задания кривых связаны с координатами и используют числовые функции. Часто удобен бескоординатный способ задания, когда для параметризации кривой используются вектор-функции. Здесь мы коротко изложим связанные с ними понятия и формулировки. Почти все доказательства опускаются.

Пусть каждому числу $t \in [a,\ b]$ по некоторому правилу поставлен в соответствие вектор $\boldsymbol{v}(t)$ трехмерного евклидова пространства. Тогда будем говорить, что на отрезке $[a,\ b]$ определена *вектор-функция* $\boldsymbol{v}(t)$. Таким образом, вектор-функция — это функция со значениями в множестве (свободных) векторов трехмерного евклидова пространства. Часто для наглядности векторы — значения вектор-функции — представляют направленными отрезками (рис. 10). Если, например, отложить все векторы $\boldsymbol{v}(t)$ из одной и той же точки O — начала от-

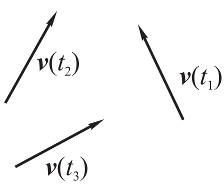


Рис. 10

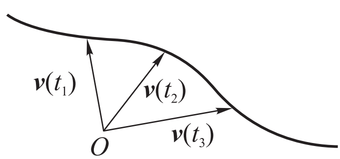


Рис. 11

счета, то их концы образуют некоторое множество точек, которое называется *годографом* вектор-функции $\boldsymbol{v}(t)$. Таким образом, годограф — это множество точек, радиус-векторы которых являются значениями вектор-функции (рис. 11).

Пусть на промежутке $[a,\ b]$ задана вектор-функция $\boldsymbol{v}(t)$. Говорят, что вектор $\boldsymbol{a}$ есть *предел* этой вектор-функции в точке $t_0 \in [a,\ b]$, если

$$\lim_{t\to t_0} |\boldsymbol{v}(t) - \boldsymbol{a}| = 0.$$

В таком случае используют запись: $\boldsymbol{a} = \lim\limits_{t \to t_0} \boldsymbol{v}(t)$. Вектор-функция $\boldsymbol{v}(t)$ называется *непрерывной в точке* t_0, если

$$\boldsymbol{v}(t_0) = \lim_{t \to t_0} \boldsymbol{v}(t).$$

Если вектор-функция $\boldsymbol{v}(t)$ непрерывна во всех точках промежутка $[a,\ b]$, то говорят, что она *непрерывна на* $[a,\ b]$.

Для вектор-функций определены те же алгебраические операции, что и для обычных векторов: это сложение, вычитание, умножение на числовую функцию, скалярное, векторное и смешанное произведения. Вводятся они поточечно:

$$\begin{gathered}
(\boldsymbol{v} + \boldsymbol{w})(t) = \boldsymbol{v}(t) + \boldsymbol{w}(t),\\
(\boldsymbol{v} - \boldsymbol{w})(t) = \boldsymbol{v}(t) - \boldsymbol{w}(t),\\
(f \cdot \boldsymbol{v})(t) = f(t) \cdot \boldsymbol{v}(t),\\
(\boldsymbol{v} \cdot \boldsymbol{w})(t) = \boldsymbol{v}(t) \cdot \boldsymbol{w}(t),\\
(\boldsymbol{v} \times \boldsymbol{w})(t) = \boldsymbol{v}(t) \times \boldsymbol{w}(t),\\
(\boldsymbol{u},\ \boldsymbol{v},\ \boldsymbol{w})(t) = (\boldsymbol{u}(t),\ \boldsymbol{v}(t),\ \boldsymbol{w}(t)).
\end{gathered}$$

Говорят, что вектор-функция $\boldsymbol{v}(t)$ *дифференцируема в точке* $t_0 \in [a,\ b]$, если при $t \to t_0$ существует предел отношения $(\boldsymbol{v}(t) - \boldsymbol{v}(t_0))/(t - t_0)$. Этот предел называется *производной* вектор-функции $\boldsymbol{v}(t)$ *в точке* t_0 и обозначается через $\boldsymbol{v}'(t)$ (рис. 12):

$$\boldsymbol{v}'(t_0) = \lim_{t \to t_0} \frac{\boldsymbol{v}(t) - \boldsymbol{v}(t_0)}{t - t_0}.$$

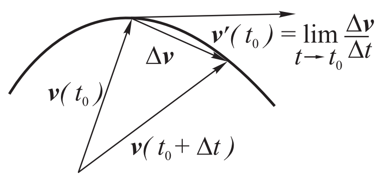


Рис. 12

Если вектор-функция дифференцируема в некоторой точке, то, очевидно, она и непрерывна в этой точке. Если вектор-функция $\boldsymbol{v}(t)$ имеет производную в каждой точке отрезка $[a,\ b]$, то говорят, что она *дифференцируема на всем отрезке* $[a,\ b]$.

Разобьем промежуток $[a,\ b]$ на n частей точками

$$a = t_0 < t_1 < t_2 < \ldots < t_n = b.$$

Пусть δ — максимальное из чисел $t_{i+1} - t_i$, при $i = 0,\ \ldots,\ n-1$. Составим интегральную сумму

$$\sigma_n = \sum_{i=0}^{n-1} \boldsymbol{v}(\tau_i)(t_{i+1} - t_i),$$

где $\tau_i \in [t_i,\ t_{i+1}]$. Будем говорить, что вектор-функция $\boldsymbol{v}(t)$ *интегрируема*, если для произвольного выбора τ_i существует предел интегральных сумм σ_n при $\delta \to 0$. Этот предел будем называть *определенным интегралом* от вектор-функции $\boldsymbol{v}(t)$ и обозначать его, как обычно:

$$\int\limits_a^b \boldsymbol{v}(t)dt.$$

Из неравенства треугольника для векторов можно вывести полезное неравенство для вектор-функций:

$$\left|\int\limits_a^b \boldsymbol{v}(t)dt\right| \leqslant \int\limits_a^b |\boldsymbol{v}(t)|dt.$$

Доказывать различные свойства вектор-функций проще всего в координатах. Фиксируем в пространстве декартову систему координат xyz с началом в некоторой точке O. Если $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ — орты координатных осей, то

$$\boldsymbol{v}(t) = v_1(t)\boldsymbol{i} + v_2(t)\boldsymbol{j} + v_3(t)\boldsymbol{k},$$

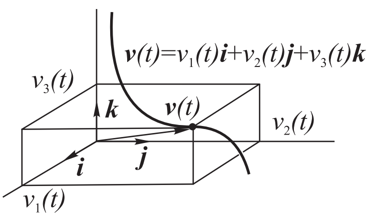


Рис. 13

где $v_1(t)$, $v_2(t)$, $v_3(t)$ — координаты вектора $\boldsymbol{v}(t)$. Функции $v_1(t)$, $v_2(t)$, $v_3(t)$ называются *координатными функциями* вектор-функции $\boldsymbol{v}(t)$ (рис. 13).

Нетрудно доказать, что вектор $\boldsymbol{a} = a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}$ является пределом функции $\boldsymbol{v}(t)$ в точке t_0 тогда и только тогда, когда его координаты a_1, a_2, a_3 являются пределами координатных функций $v_1(t)$, $v_2(t)$, $v_3(t)$. Из этого легко следует, что если $f(t) \to \alpha$, $\boldsymbol{u}(t) \to \boldsymbol{a}$, $\boldsymbol{v}(t) \to \boldsymbol{b}$, $\boldsymbol{w}(t) \to \boldsymbol{c}$ при $t \to t_0$, то

$$\begin{gathered}
\boldsymbol{v}(t) + \boldsymbol{w}(t) \to \boldsymbol{b} + \boldsymbol{c}, \qquad \boldsymbol{v}(t) - \boldsymbol{w}(t) \to \boldsymbol{b} - \boldsymbol{c},\\
f(t) \cdot \boldsymbol{v}(t) \to \alpha\boldsymbol{b}, \qquad \boldsymbol{v}(t) \cdot \boldsymbol{w}(t) \to \boldsymbol{b} \cdot \boldsymbol{c},\\
\boldsymbol{v}(t) \times \boldsymbol{w}(t) \to \boldsymbol{b} \times \boldsymbol{c}, \quad (\boldsymbol{u}(t),\ \boldsymbol{v}(t),\ \boldsymbol{w}(t)) \to (\boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{c}).
\end{gathered}$$

Кроме того, ясно, что вектор-функция $\boldsymbol{v}(t)$ непрерывна в точке t_0 (на всем отрезке $[a,\ b]$) тогда и только тогда, когда в точке t_0 (на отрезке

$[a,\ b]$) непрерывны ее координатные функции. Из этого следует, что если функции $f(t)$, $\boldsymbol{u}(t)$, $\boldsymbol{v}(t)$, $\boldsymbol{w}(t)$ непрерывны в точке t_0 (на отрезке $[a,\ b]$), то вместе с ними непрерывны функции $\boldsymbol{u}+\boldsymbol{v}$, $\boldsymbol{u}-\boldsymbol{v}$, $f\cdot\boldsymbol{u}$, $\boldsymbol{u}\cdot\boldsymbol{v}$, $\boldsymbol{u}\times\boldsymbol{v}$, $(\boldsymbol{u},\ \boldsymbol{v},\ \boldsymbol{w})$.

Нетрудно доказать, что вектор-функция $\boldsymbol{v}(t)$ дифференцируема в точке $t_0\in[a,\ b]$ тогда и только тогда, когда в этой точке дифференцируемы координатные функции $v_1(t)$, $v_2(t)$, $v_3(t)$. При этом

$$\boldsymbol{v}'(t_0)=v_1'(t_0)\boldsymbol{i}+v_2'(t_0)\boldsymbol{j}+v_3'(t_0)\boldsymbol{k}.$$

Если $\boldsymbol{v}(t)$ дифференцируема на всем отрезке $[a,\ b]$, то ее координатные функции также дифференцируемы. Верно и обратное утверждение. При этом если $f(t)$, $\boldsymbol{u}(t)$, $\boldsymbol{v}(t)$, $\boldsymbol{w}(t)$ дифференцируемы на $[a,\ b]$, то

$$\begin{gathered}(\boldsymbol{v}+\boldsymbol{w})'=\boldsymbol{v}'+\boldsymbol{w}',\quad(\boldsymbol{v}-\boldsymbol{w})'=\boldsymbol{v}'-\boldsymbol{w}',\\(f\cdot\boldsymbol{v})'=f'\boldsymbol{v}+f\boldsymbol{v}',\quad(\boldsymbol{v}\cdot\boldsymbol{w})'=\boldsymbol{v}'\cdot\boldsymbol{w}+\boldsymbol{v}\cdot\boldsymbol{w}'\\(\boldsymbol{v}\times\boldsymbol{w})'=\boldsymbol{v}'\times\boldsymbol{w}+\boldsymbol{v}\times\boldsymbol{w}',\\(\boldsymbol{u},\ \boldsymbol{v},\ \boldsymbol{w})'=(\boldsymbol{u}',\ \boldsymbol{v},\ \boldsymbol{w})+(\boldsymbol{u},\ \boldsymbol{v}',\ \boldsymbol{w})+(\boldsymbol{u},\ \boldsymbol{v},\ \boldsymbol{w}').\end{gathered}$$

Аналогично обстоит дело с интегрируемостью. Вектор-функция $\boldsymbol{v}(t)$ интегрируема тогда и только тогда, когда интегрируемы ее координатные функции v_1, v_2, v_3. При этом

$$\int\limits_a^b\boldsymbol{v}(t)dt=\boldsymbol{i}\int\limits_a^b v_1(t)dt+\boldsymbol{j}\int\limits_a^b v_2(t)dt+\boldsymbol{k}\int\limits_a^b v_3(t)dt.$$

Ясно теперь, что всякая непрерывная вектор-функция интегрируема.

Если функция $\boldsymbol{v}(t)$ имеет непрерывную производную, то справедлива формула Ньютона—Лейбница

$$\boldsymbol{v}(b)-\boldsymbol{v}(a)=\int\limits_a^b\boldsymbol{v}'(t)dt.$$

Векторные уравнения кривых. Как мы скоро убедимся, вектор-функции очень удобны для описания и исследования кривых. Если $F\colon[a,\ b]\to R^3$ — параметризация некоторой кривой C, то ей соответствует вектор-функция $\boldsymbol{f}$, определенная по формуле

$$\boldsymbol{f}(t)=\overrightarrow{OF(t)}.$$

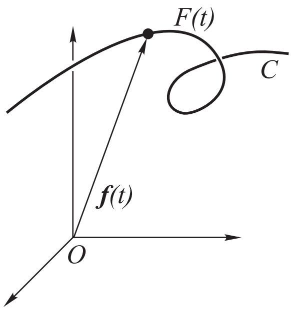


Рис. 14

Функция F однозначно восстанавливается по вектор-функции $\boldsymbol{f}$. Координатные функции у $\boldsymbol{f}$ и F, как видно из определения, совпадают. Вектор-функция $\boldsymbol{f}$ называется *векторной параметризацией* кривой C (рис. 14). Если радиус-вектор $\overrightarrow{OF(t)}$ обозначать через $\boldsymbol{r}$, то равенство

$$\boldsymbol{r} = \boldsymbol{f}(t)$$

называется *векторным уравнением* кривой C. Кривую C при этом можно рассматривать как годограф вектор-функции $\boldsymbol{f}(t)$.

Вектор-функция $\boldsymbol{f}$ непрерывна в силу непрерывности отображения F. Если кривая C гладкая, а F — ее регулярная параметризация, то функция $\boldsymbol{f}$ непрерывно дифференцируема на отрезке $[a,\ b]$, причем ее производная нигде не обращается в нуль:

$$\boldsymbol{f}'(t) \neq \boldsymbol{0}.$$

Закончим параграф одной леммой о вектор-функциях, которой мы часто будем пользоваться при изучении кривых.

Лемма. *Пусть $\boldsymbol{v}(t)$ — вектор-функция, дифференцируемая всюду на отрезке $[a,\ b]$. Тогда для того, чтобы ее модуль $|v(t)|$ был постоянной функцией, необходимо и достаточно, чтобы $\boldsymbol{v}(t)$ была всюду ортогональна своей производной $\boldsymbol{v}'(t) : \boldsymbol{v}(t) \cdot \boldsymbol{v}'(t) \equiv \boldsymbol{0}$.*

Доказательство. Ясно, что $|v(t)|$ — постоянная функция тогда и только тогда, когда $\boldsymbol{v}^2(t)$ — постоянная функция. Но производная этой функции как раз и совпадает с $2\boldsymbol{v}(t) \cdot \boldsymbol{v}'(t) \equiv \boldsymbol{0}$. Осталось заметить, что равенство нулю производной равносильно постоянству самой функции. □

В дальнейшем все параметризации у нас будут только векторными. При этом мы позволим себе отождествлять точку с ее радиус-вектором и писать $\boldsymbol{f}(t) = P$ вместо $\boldsymbol{f}(t) = \overrightarrow{OP}$.

§ 3. Касательная кривой

Пусть C — гладкая элементарная кривая, а $\boldsymbol{f}(t)$ вектор-функция, задающая ее регулярную параметризацию. Если $P = \boldsymbol{f}(t_0)$ — точка кривой, то вектор $\boldsymbol{f}'(t_0)$ называется *касательным вектором кривой* C в точке P (рис. 15).

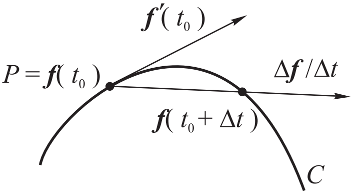


Рис. 15

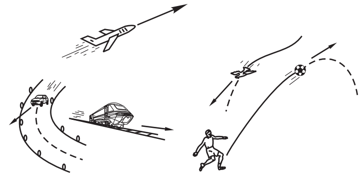


Рис. 16

Если вектор-функция $\boldsymbol{f}(t)$ описывает перемещение материальной точки вдоль кривой C, то вектор $\boldsymbol{f}\,'(t_0)$ будет *вектором скорости* этой точки в момент времени t_0 (рис. 16).

Касательные векторы в одной и той же точке, соответствующие различным параметризациям, коллинеарны и, значит, могут отличаться только множителем.

Действительно, если $\boldsymbol{g}(\tau) = \boldsymbol{f}(\varphi(\tau))$ — другая параметризация той же кривой, причем $t_0 = \varphi(\tau_0)$, то вектор $\boldsymbol{g}'(\tau_0) = \boldsymbol{f}\,'(\varphi(\tau_0)) \cdot \varphi'(\tau_0)$ очевидно коллинеарен вектору $\boldsymbol{f}\,'(t_0) = \boldsymbol{f}\,'(\varphi(\tau_0))$. $\square$

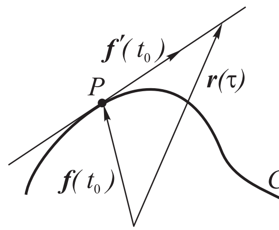


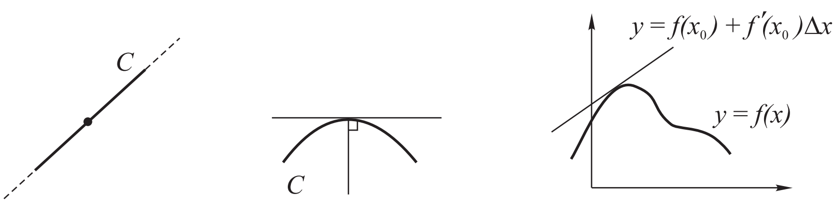
Рис. 17

Прямая, проходящая через точку P в направлении касательного вектора $\boldsymbol{f}\,'(t_0)$, называется *касательной прямой в точке* P (рис. 17). Параметрическое уравнение касательной прямой имеет вид

$$\boldsymbol{r}(\tau) = \boldsymbol{f}(t_0) + \tau \boldsymbol{f}\,'(t_0).$$

Примеры. 1. Касательная прямая к отрезку в любой точке совпадает с прямой, на которой лежит этот отрезок (рис. 18, *а*).

2. Касательная к дуге окружности совпадает с обычной касательной к самой окружности (в той же точке, рис. 18, *б*).



а *б* *в*

Рис. 18

3. Касательная к графику гладкой функции $f(x)$ в точке $(x_0,\ f(x_0))$ имеет уравнение (рис. 18, *в*)

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Наше определение касательной не очень геометрично, хотя удобно на практике. За основу другого определения можно взять следующее важное свойство касательной. Пусть Q — точка кривой C, близкая к точке P.

Теорема. *При стремлении точки Q кривой C к точке P предел отношения расстояния δ от точки Q до касательной прямой в точке P к расстоянию d от Q до P равен нулю*:

$$\lim_{Q\to P} \frac{\delta}{d} = 0.$$

Касательная является единственной прямой, обладающей этим свойством (рис. 19).

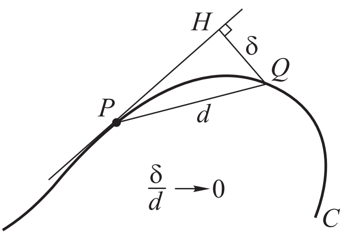


Рис. 19

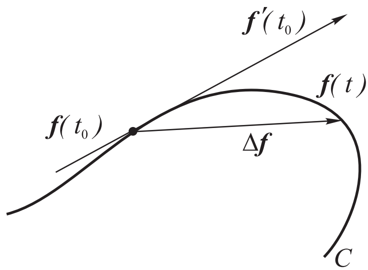


Рис. 20

Доказательство. Пусть $Q = \boldsymbol{f}(t)$. Положим $\Delta t = t - t_0$, $\Delta\boldsymbol{f} = \overrightarrow{PQ} = \boldsymbol{f}(t) - \boldsymbol{f}(t_0)$ (рис. 20). Расстояние d от точки Q до точки P равняется $|\Delta\boldsymbol{f}|$, а расстояние δ от точки Q до касательной в точке P равняется $|\Delta\boldsymbol{f} \times \boldsymbol{f}'(t_0)| : |\boldsymbol{f}'(t_0)|$. Наконец,

$$\lim_{Q\to P} \frac{\delta}{d} = \lim_{t\to t_0} \frac{\delta/\Delta t}{d/\Delta t} =$$
$$= \frac{\lim\limits_{t\to t_0} \frac{\delta}{\Delta t}}{\lim\limits_{t\to t_0} \frac{d}{\Delta t}} = \frac{\lim\limits_{t\to t_0} \left|\frac{\Delta\boldsymbol{f}}{\Delta t} \times \boldsymbol{f}'(t_0)\right|}{\lim\limits_{t\to t_0} \left|\frac{\Delta\boldsymbol{f}}{\Delta t}\right| \cdot |\boldsymbol{f}'(t_0)|} = \frac{|\boldsymbol{f}'(t_0) \times \boldsymbol{f}'(t_0)|}{|\boldsymbol{f}'(t_0)|^2} = 0.$$

Первая часть теоремы доказана.

Пусть теперь l — произвольная прямая, проходящая через точку P, а $\boldsymbol{b}$ — ее направляющий вектор. Обозначим через δ_1 расстояние от точки Q до прямой l (рис. 21). Воспроизводя для прямой l предыдущие выкладки, получаем, что

$$\lim_{Q\to P} \frac{\delta_1}{d} = \frac{|\boldsymbol{f}\,'(t_0) \times \boldsymbol{b}|}{|\boldsymbol{f}\,'(t_0)| \cdot |\boldsymbol{b}|}.$$

Осталось заметить, что этот предел равен нулю в том и только в том случае, когда векторы $\boldsymbol{f}\,'(t_0)$ и $\boldsymbol{b}$ коллинеарны, т. е. когда прямая l является касательной. Теорема доказана. □

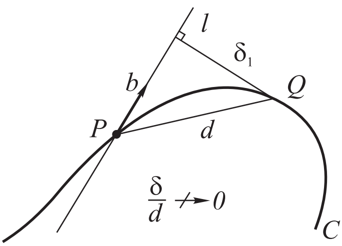


Рис. 21

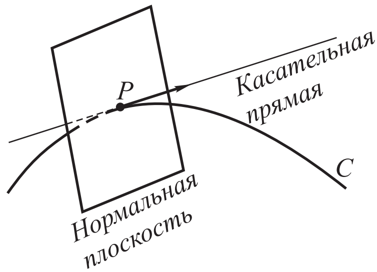


Рис. 22

Плоскость, содержащая точку P и ортогональная касательной прямой, называется *нормальной плоскостью* кривой C в точке P (рис. 22).

§ 4. Длина кривой

Пусть элементарная кривая C параметризована вектор-функцией $\boldsymbol{f}\colon [a,\ b] \to \mathbf{R}^3$. На отрезке $[a,\ b]$ выберем $n-1$ точек, разбивающих его на n частей: $a < t_1 < t_2 < \ldots < t_{n-1} < b$. Для удобства обозначений положим $t_0 = a$, $t_n = b$. Ломаная с вершинами в точках $\boldsymbol{f}(t_0)$, $\boldsymbol{f}(t_1)$, $\ldots$, $\boldsymbol{f}(t_n)$ называется *вписанной* в кривую C (рис. 23).

Интуитивно кажется ясным, что длина кривой C (если о ней вообще можно говорить) не должна сильно отличаться от длины вписанной ломаной при условии, что у ломаной достаточно много звеньев и все они достаточно малы (рис. 24). Это приводит нас к следующему определению. *Длиной* кривой C называется предел, к которому стремится длина вписанных в нее ломаных при неограниченном возрастании числа звеньев ломаной и неограниченном убывании их длин.

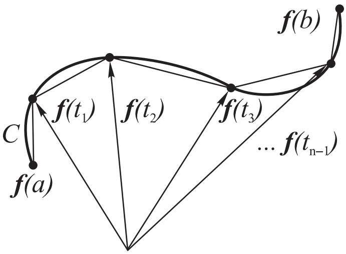


Рис. 23

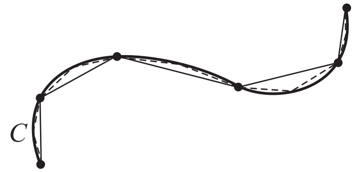


Рис. 24

Отметим, что условие убывания длин звеньев можно заменить другим, равносильным ему условием: надо потребовать, чтобы

$$\max_{i=0,\,\ldots,\,n-1}(t_{i+1}-t_i)\to 0.$$

Нетрудно видеть, что этот предел (конечный или бесконечный) всегда существует: это есть не что иное, как супремум длин ломаных, вписанных в кривую.

Кривая C называется *спрямляемой*, если ее длина конечна.

Теорема. *Всякая элементарная гладкая кривая C спрямляема. Ее длина S может быть найдена по формуле*

$$S=\int\limits_a^b|\boldsymbol{f}'(t)|dt,$$

где $\boldsymbol{f}$: $[a,\ b]\to\boldsymbol{R}^3$ — *произвольная регулярная параметризация кривой C.*

Доказательство. Оценим разность между длиной ломаной и интегралом модуля производной. Для этого введем обозначения:

$$\Delta_i\boldsymbol{f}=\boldsymbol{f}(t_i)-\boldsymbol{f}(t_{i-1}),\quad \boldsymbol{f}'_i=\boldsymbol{f}'(t_i),\quad \Delta_i t=t_i-t_{i-1}.$$

Тогда имеем

$$\left|\sum_{i=1}^n|\Delta_i\boldsymbol{f}|-\int\limits_a^b|\boldsymbol{f}'(t)|dt\right|\leqslant\left|\sum_{i=1}^n|\Delta_i\boldsymbol{f}|-\sum_{i=1}^n|\boldsymbol{f}'_i|\Delta_i t\right|+$$

$$+\left|\sum_{i=1}^n|\boldsymbol{f}'_i|\Delta_i t-\int\limits_a^b|\boldsymbol{f}'(t)|dt\right|=I+II.$$

Второе слагаемое по мере уменьшения звеньев ломаной стремится к нулю по определению интеграла. Осталось доказать, что к нулю стремится и первое слагаемое. Преобразовывая, получаем неравенство:

$$I = \left| \sum_{i=1}^{n} (|\Delta_i \boldsymbol{f}| - |\boldsymbol{f}\,'_i|\Delta_i t) \right| \leqslant \sum_{i=1}^{n} |\Delta_i \boldsymbol{f} - \boldsymbol{f}\,'_i \Delta_i t|.$$

Оценим отдельно каждое слагаемое в этой сумме:

$$|\Delta_i \boldsymbol{f} - \boldsymbol{f}\,'_i \Delta_i t| = \left| \int\limits_{t_{i-1}}^{t_i} \boldsymbol{f}\,'(t)dt - \int\limits_{t_{i-1}}^{t_i} \boldsymbol{f}\,'_i dt \right| =$$
$$= \left| \int\limits_{t_{i-1}}^{t_i} (\boldsymbol{f}(t) - \boldsymbol{f}\,'_i)dt \right| \leqslant \int\limits_{t_{i-1}}^{t_i} |\boldsymbol{f}\,'(t) - \boldsymbol{f}\,'_i|dt.$$

Так как функция $f'(t)$ непрерывна на замкнутом интервале $[a,\ b]$, то она и равномерно непрерывна. Это значит, что для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что если $x,\, y \in [a,\ b]$ и $|x-y| < \delta$, то $|\boldsymbol{f}(x) - \boldsymbol{f}(y)| < \varepsilon$. Поэтому если все отрезки $[t_{i-1} t_i]$ по длине меньше δ, то $|\boldsymbol{f}\,'(t) - \boldsymbol{f}\,'_i| < \varepsilon$ для любого $t \in [t_{i-1},\ t_i]$, и, следовательно,

$$|\Delta_i \boldsymbol{f} - \boldsymbol{f}\,'\Delta_i t| \leqslant \int\limits_{t_{i-1}}^{t_i} \varepsilon \cdot dt = \varepsilon \cdot \Delta_i t.$$

Окончательно получаем, что для достаточно мелких разбиений

$$I \leqslant \sum_{i=1}^{n} (\varepsilon \cdot \Delta_i t) = \varepsilon \cdot \sum_{i=1}^{n} \Delta_i t = \varepsilon(t_n - t_0).$$

Отсюда в силу произвольности выбора числа ε следует, что первое слагаемое стремится к нулю. $\square$

Примеры. 1. Если $f_1,\ f_2,\ f_3$ — координатные функции вектор-функции $\boldsymbol{f}$, то формула длины кривой примет вид

$$S = \int\limits_a^b \sqrt{(f_1'(t))^2 + (f_2'(t))^2 + (f_3'(t))^2}\, dt$$

или, короче,

$$S = \int\limits_a^b \sqrt{f_1'^2 + f_2'^2 + f_3'^2}\, dt.$$

2. Если кривая C плоская и явно задана уравнением $y = f(x)$, то, подставляя в предыдущую формулу $t = x$, $f_1(x) = x$, $f_2(x) = f(x)$, $f_3(x) = 0$, получаем

$$S = \int\limits_a^b \sqrt{1 + f'^2}\, dx.$$

Естественная параметризация. Длину дуги можно использовать для введения одной очень удобной параметризации кривой C. Зададим на промежутке $[a,\ b]$ функцию $\psi(t)$ по формуле

$$\psi(t) = \int\limits_a^t |\boldsymbol{f}\,'(\tau)| d\tau.$$

Ясно, что $\psi(t)$ равняется длине дуги с началом в $\boldsymbol{f}(a)$ и концом $\boldsymbol{f}(t)$. Функция $\psi(t)$ гладкая: очевидно, $\psi'(t) = |\boldsymbol{f}\,'(t)|$. Она монотонно возрастает (поскольку ее производная $|\boldsymbol{f}\,'(t)|$ положительна) и отображает отрезок $[a,\ b]$ в отрезок $[0,\ S]$.

Рассмотрим обратную функцию $\varphi = \psi^{-1}\colon [0,\ S] \to [a,\ b]$,

$$\varphi(s) = \psi^{-1}(s),$$

и параметризацию $\boldsymbol{g}(s)$ кривой C, получающуюся из параметризации $\boldsymbol{f}(t)$ при замене параметра $t = \varphi(s)$:

$$\boldsymbol{g}(s) = \boldsymbol{f}(\varphi(s)).$$

Такая параметризация называется *естественной параметризацией кривой* C (рис. 25). (Другие названия: натуральная параметризация или параметризация длиной дуги.) Параметром в этой параметризации служит s — длина отрезка кривой. Это и есть так называемый *естественный параметр.* Параметризация $\boldsymbol{f}(t)$ получается из $\boldsymbol{g}(s)$ при обратной замене параметра $s = \psi(t)$.

Важное свойство естественной параметризации состоит в том, что касательный вектор при такой параметризации имеет единичную длину (рис. 26).

$$|\boldsymbol{g}'(s)| \equiv 1.$$

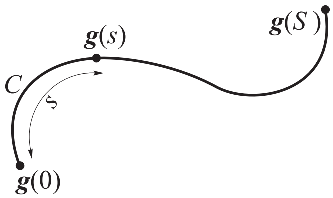


Рис. 25

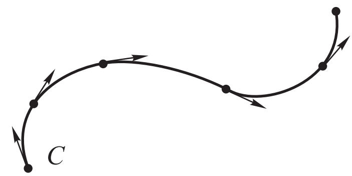


Рис. 26

Действительно, поскольку

$$\varphi'(s) = (\psi^{-1})'(s) = \frac{1}{\psi'(\varphi(s))} = \frac{1}{|\boldsymbol{f}\,'(\varphi(s))|},$$

то

$$|\boldsymbol{g}'(s)| = |\boldsymbol{f}\,'(\varphi(s))\varphi'(s)| = \left|\frac{\boldsymbol{f}\,'(\varphi(s))}{|\boldsymbol{f}\,'(\varphi(s))|}\right| \equiv 1.$$

Этим своим свойством и тем, что $\boldsymbol{g}(0)$ совпадает с началом кривой, естественная параметризация определена однозначно.

Пример. Уравнения $x = \cos t$, $y = \sin t$ при $t \in [o, \pi]$ задают естественную параметризацию верхней единичной полуокружности.

В дальнейшем мы будем обозначать вектор $\boldsymbol{g}\,'(s)$ через $\boldsymbol{t}$ и называть его *единичным вектором касательной*, или *единичным касательным вектором*, кривой в точке $\boldsymbol{g}(s)$.

§ 5. Кривизна кривой. Соприкасающаяся плоскость

Пусть C — гладкая кривая, а $\boldsymbol{g}(s)$ — ее естественная параметризация. Рассмотрим произвольную точку $P = \boldsymbol{g}(s_0)$ на кривой C и вектор $\boldsymbol{g}\,''(s_0)$ второй производной функции $\boldsymbol{g}$ в этой точке. Этот вектор называется *вектором кривизны* кривой C в точке P и обозначается через $\boldsymbol{k}$ (рис. 27):

$$\boldsymbol{k} = \boldsymbol{g}\,''(s_0).$$

Его длина $k = |\boldsymbol{k}|$ называется *кривизной* кривой C в точке P.

Поскольку $|\boldsymbol{g}\,'(s)| \equiv 1$, то вектор $\boldsymbol{k} = \boldsymbol{g}\,''(s_0)$ ортогонален вектору $\boldsymbol{g}\,'(s_0)$ и, следовательно, лежит в нормальной плоскости. Если $k \neq 0$, то $\boldsymbol{k} \neq \boldsymbol{0}$ и прямая, проходящая через точку P в направлении вектора $\boldsymbol{k}$, называется *главной нормалью* кривой C в точке P. В этом случае единичный вектор $\boldsymbol{n} = \boldsymbol{k}/k$ называется *единичным вектором главной нормали* (рис. 28).

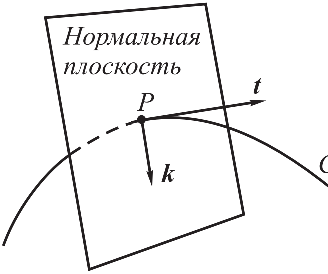


Рис. 27

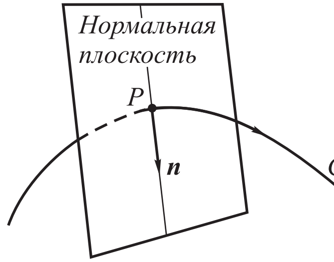


Рис. 28

Таким образом, имеет место равенство

$$\frac{d\boldsymbol{t}}{ds} = k \cdot \boldsymbol{n}.$$

Это так называемая «первая формула Френе».

Кривизна позволяет определить, насколько данная кривая отличается от прямой. Так, кривизна прямой во всех точках равна нулю. Верно и обратное: если кривизна кривой равна нулю во всех точках, то кривая является отрезком прямой. (Докажите.)

Пример. Кривизна дуги окружности радиуса R во всех точках равняется $1/R$. (Проверьте.)

Величина, обратная кривизне, называется *радиусом кривизны* кривой в данной точке.

Пусть $\boldsymbol{f}(t)$ — произвольная параметризация кривой C. Пусть она связана с естественной параметризацией $\boldsymbol{g}(s)$ при помощи замены параметра: $s = \psi(t) = \int\limits_0^t |\boldsymbol{f}'(\tau)|d\tau$, причем $s_0 = \psi(t_0)$. Вектор $\boldsymbol{f}''(t_0)$ может не быть коллинеарным вектору кривизны $\boldsymbol{k} = \boldsymbol{g}''(s_0)$, но его проекция $[\boldsymbol{f}''(t_0)]^{\perp}$ на нормальную плоскость коллинеарна вектору $\boldsymbol{k}$. Действительно, $\boldsymbol{f}'(t_0) = \boldsymbol{g}'(s_0) \cdot \psi'(t_0)$. Следовательно,

$$\boldsymbol{f}''(t_0) = \boldsymbol{g}''(s_0) \cdot (\psi'(t_0))^2 + \boldsymbol{g}'(s_0) \cdot \psi''(t_0).$$

Так как первое слагаемое лежит в нормальной плоскости, а второе — на касательной прямой, то первое слагаемое и равняется нормальной составляющей вектора $\boldsymbol{f}''(t_0)$ (рис. 29):

$$(\boldsymbol{f}''(t_0))^{\perp} = |\boldsymbol{f}'(t_0)|^2 \cdot \boldsymbol{k}.$$

(Механический смысл этой формулы заключается в том, что *нормальное ускорение* материальной точки, движущейся по кривой C, зависит только от скорости: оно прямо пропорционально квадрату скорости, а коэффициентом пропорциональности служит вектор кривизны.)

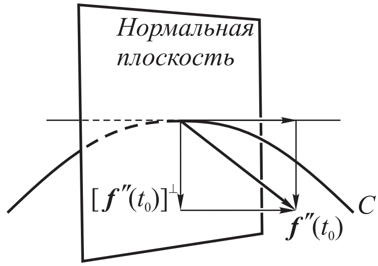


Рис. 29

В нашем определении кривизна k является функцией точки P, но часто удобно считать, что она зависит от естественного параметра s или от того же параметра t, что и функция $\boldsymbol{f}$. Найдем выражение кривизны через первую и вторую производные функции $\boldsymbol{f}$. Ясно, что $k(t_0) = \left|(\boldsymbol{f}''(t_0))^{\perp}/|\boldsymbol{f}'(t_0)|^2\right|$. Обозначим через α угол между векторами $\boldsymbol{f}'(t_0)$ и $\boldsymbol{f}''(t_0)$. Тогда

$$|(\boldsymbol{f}''(t_0))^{\perp}| = |\boldsymbol{f}''(t_0)| \cdot \sin\alpha = \frac{|\boldsymbol{f}''(t_0) \times \boldsymbol{f}'(t_0)|}{|\boldsymbol{f}'(t_0)|}.$$

Окончательно получаем

$$k(t_0) = \frac{|\boldsymbol{f}''(t_0) \times \boldsymbol{f}'(t_0)|}{|\boldsymbol{f}'(t_0)|^3},$$

или, короче,

$$k = \frac{|\boldsymbol{f}'' \times \boldsymbol{f}'|}{|\boldsymbol{f}'|^3}.$$

Пример. Кривизна плоской кривой, заданной уравнением в явном виде $y = f(x)$, вычисляется по формуле $k = |f''|/(1 + f'^2)^{3/2}$. Отметим, что в тех точках, где первая производная у $f(x)$ обращается в нуль, кривизна просто равна модулю второй производной: $k = |f''|$.

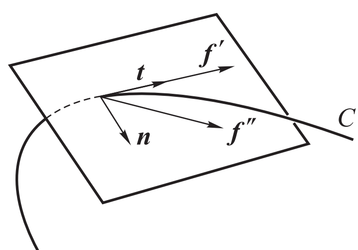


Рис. 30

Пусть в точке P кривая C имеет ненулевую кривизну. Как мы видели, векторы $\boldsymbol{f}'(t_0)$ и $\boldsymbol{f}''(t_0)$ всегда

лежат в плоскости, натянутой на векторы $\boldsymbol{t}$ и $\boldsymbol{n}$. Эта плоскость называется *соприкасающейся плоскостью* кривой C в точке P (рис. 30, 31). В качестве нормального вектора к ней можно взять единичный

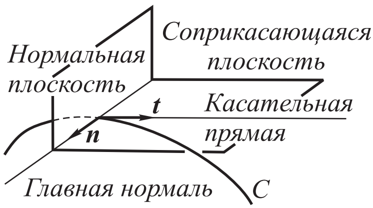


Рис. 31

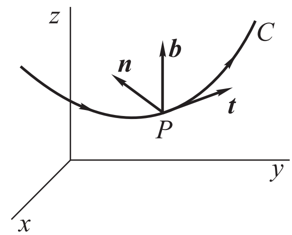


Рис. 32

вектор $\boldsymbol{b} = \boldsymbol{t} \times \boldsymbol{n}$, который называется *вектором бинормали* в точке P. Нетрудно видеть, что в каждой точке кривой C, где кривизна отлична от нуля, векторы $\boldsymbol{t}$, $\boldsymbol{n}$, $\boldsymbol{b}$ образуют правую тройку взаимно ортогональных векторов. Эта тройка векторов называется *базисом Френе* (рис. 32). (Другие названия: репер Френе, трехвекторник Френе, сопровождающий трехвекторник и т. п.) Очевидно, что

$$\boldsymbol{b} = \frac{\boldsymbol{f}'(t_0) \times \boldsymbol{f}''(t_0)}{|\boldsymbol{f}'(t_0) \times \boldsymbol{f}''(t_0)|}.$$

Соприкасающаяся плоскость плоской кривой всегда совпадает с той плоскостью, в которой лежит кривая.

Соприкасающаяся плоскость обладает одним важным свойством, которое можно взять за основу при геометрическом ее определении:

Теорема. *Пусть Q — точка кривой C, близкая к точке P. При стремлении точки Q к точке P отношение расстояния δ от точки Q до соприкасающейся плоскости в точке P к квадрату расстояния d от Q до P стремится к нулю*:

$$\lim_{Q \to P} \frac{\delta}{d^2} = 0.$$

Если кривизна кривой C в точке P отлична от нуля, то единственной плоскостью, обладающей таким свойством, будет соприкасающаяся плоскость (рис. 33).

(Сформулированная теорема означает, что в каждой точке кривая с точностью до величин второго порядка малости приближается плоской кривой — своей проекцией на соприкасающуюся плоскость в этой точке.) Этот факт мы оставим без доказательства, тем более что нигде не будем его использовать. □

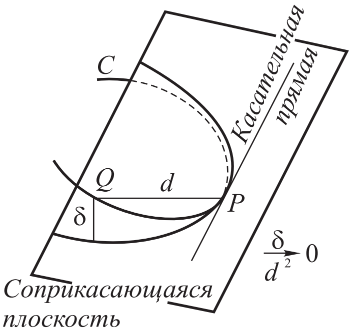


Рис. 33

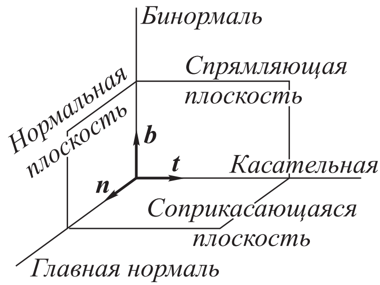


Рис. 34

Плоскость, проходящая через точку P и содержащая векторы $\boldsymbol{b}$ и $\boldsymbol{t}$, называется *спрямляющей плоскостью* кривой C в точке P (рис. 34).

§ 6. Кручение кривой. Формулы Френе

Пусть C — гладкая кривая, а $\boldsymbol{g}(s)$ — ее естественная параметризация. В этом и двух следующих параграфах мы будем предполагать, что во всех своих точках кривая C имеет ненулевую кривизну. В частности, в любой ее точке $P = \boldsymbol{g}(s_0)$ определены соприкасающаяся плоскость и вектор бинормали

$$\boldsymbol{b}(s_0) = \boldsymbol{t}(s_0) \times \boldsymbol{n}(s_0) = \frac{\boldsymbol{g}\,'(s_0) \times \boldsymbol{g}\,''(s_0)}{k(s_0)}.$$

Рассмотрим вектор $\boldsymbol{b}'(s_0)$ первой производной от функции $\boldsymbol{b}(s)$. Он ортогонален вектору $\boldsymbol{t}(s_0)$, поскольку

$$\boldsymbol{b}'(s) = [\boldsymbol{t}(s) \times \boldsymbol{n}(s)]' = \boldsymbol{t}'(s) \times \boldsymbol{n}(s) + \boldsymbol{t}(s) \times \boldsymbol{n}'(s) = \boldsymbol{t}(s) \times \boldsymbol{n}'(s).$$

(Поясним третье равенство. Первое слагаемое в левой его части равно нулю, поскольку векторы $\boldsymbol{t}'$ и $\boldsymbol{n}$ коллинеарны (первая формула

Френе).) Далее, вектор $\boldsymbol{b}'(s_0)$ ортогонален вектору $\boldsymbol{b}(s_0)$ ввиду постоянства модуля вектор-функции $\boldsymbol{b}(s)$. Следовательно, вектор $\boldsymbol{b}'(s_0)$ коллинеарен вектору $\boldsymbol{n}(s_0)$. Поэтому выполняется равенство вида

$$\boldsymbol{b}'(s_0) = -\varkappa \cdot \boldsymbol{n}(s_0).$$

Число $\varkappa$ называется кручением кривой C в точке P. Выписанное равенство носит название «третьей формулы Френе». *Абсолютным кручением* $|\varkappa|$ в точке P называется абсолютная величина вектора $\boldsymbol{b}'(s_0)$:

$$|\varkappa| = |\boldsymbol{b}'(s_0)|.$$

Кручение характеризует отличие пространственной кривой от плоской, поскольку, очевидно, кручение плоской кривой в каждой точке равно нулю. Нетрудно видеть, что если кручение кривой в каждой точке равно нулю, то эта кривая лежит в некоторой плоскости (рис. 35).

Абсолютное кручение можно определить более геометрически. Если Q — близкая к P точка кривой C, а θ — угол между соприкасающимися плоскостями кривой C в точках P и Q, то *при стремлении*

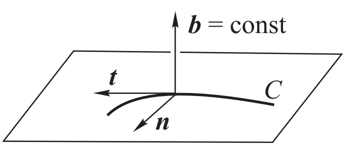


Рис. 35

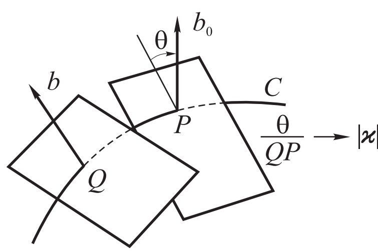


Рис. 36

точки Q к точке P отношение угла θ к расстоянию между Q и P стремится к определенному пределу, который и равен абсолютному кручению кривой C в точке P (рис. 36): $\lim\limits_{Q\to P} \theta/QP = |\varkappa|$.

Кручение допускает и несложное кинематическое истолкование. Представим себе, что некоторая плоскость перемещается в пространстве, причем ее фиксированная точка с единичной скоростью движется по кривой, фиксированная прямая в каждый момент времени касается кривой в этой точке, а сама плоскость все время является соприкасающейся плоскостью кривой. Тогда такое перемещение будет результатом поступательного движения и двух вращений — вращения этой

плоскости вокруг бинормали и ее вращения вокруг касательной. Угловая скорость первого вращения равна кривизне кривой, а второго — абсолютному кручению кривой в точке соприкосновения. Знак кручения связан с направлением вращения: в случае, когда вращение происходит против часовой стрелки, если смотреть из конца касательного вектора, то это плюс, а если по часовой стрелке — то минус (рис. 37).

Полученные нами ранее выражения для производных по естественному параметру вектор-функций $\boldsymbol{t}$ и $\boldsymbol{b}$ представляют собой две из трех *формул Френе* (называемых также формулами Френе — Серре). Эти формулы дают разложение по базису Френе $\{\boldsymbol{t},\ \boldsymbol{n},\ \boldsymbol{b}\}$ производных от входящих в него векторов:

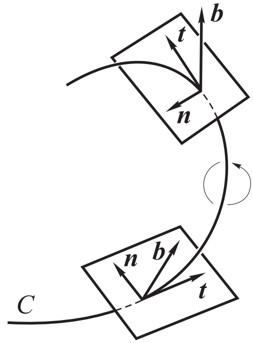


Рис. 37

$$\begin{aligned} \boldsymbol{t}'(s) &= \qquad\quad k(s)\boldsymbol{n}(s), \\ \boldsymbol{n}'(s) &= -k(s)\boldsymbol{t}(s) + \varkappa(s)\boldsymbol{b}(s), \\ \boldsymbol{b}'(s) &= \qquad -\varkappa(s)\boldsymbol{n}(s). \end{aligned}$$

Первая и третья формулы нам уже известны. Вторая формула из них следует:

$$\boldsymbol{n}' = (\boldsymbol{b} \times \boldsymbol{t})' = \boldsymbol{b}' \times \boldsymbol{t} + \boldsymbol{b} \times \boldsymbol{t}' = -\varkappa(\boldsymbol{n} \times \boldsymbol{t}) + k(\boldsymbol{b} \times \boldsymbol{n}) = \varkappa\boldsymbol{b} - k\boldsymbol{t}.$$

Иногда, имея в виду уравнение кривой $\boldsymbol{r} = \boldsymbol{g}(s)$, к числу формул Френе относят и формулу $\boldsymbol{t} = \boldsymbol{r}'$.

§ 7. Вычисление кручения

Пусть по-прежнему $\boldsymbol{g}(s)$ — естественная параметризация гладкой кривой C, имеющей в каждой точке ненулевую кривизну.

Для вычисления кручения кривой C воспользуемся формулами Френе:

$$\boldsymbol{g}\,'(s) = \boldsymbol{t}, \quad \boldsymbol{g}\,''(s) = k\boldsymbol{n},$$

$$\boldsymbol{g}\,'''(s) = (k\boldsymbol{n})' = k'\boldsymbol{n} + k\boldsymbol{n}' = k'\boldsymbol{n} + k(-k\boldsymbol{t} + \varkappa\boldsymbol{b}) = -k^2\boldsymbol{t} + k'\boldsymbol{n} + k\varkappa\boldsymbol{b}.$$

Рассмотрим теперь смешанное произведение $(\boldsymbol{g}\,', \boldsymbol{g}\,'', \boldsymbol{g}\,''')$. Так как смешанное произведение не изменится, если прибавить к одному сомножителю или вычесть из него линейную комбинацию двух других,

то

$$(\boldsymbol{g}'(s),\ \boldsymbol{g}''(s),\ \boldsymbol{g}'''(s)) = (\boldsymbol{t},\ k\boldsymbol{n},\ -k^2\boldsymbol{t} + k'\boldsymbol{n} + k\varkappa\boldsymbol{b}) =$$
$$= (\boldsymbol{t},\ k\boldsymbol{n},\ k\varkappa\boldsymbol{b}) = k^2(s)\cdot\varkappa(s).$$

Окончательно получаем

$$\varkappa = \frac{(\boldsymbol{g}',\ \boldsymbol{g}'',\ \boldsymbol{g}''')}{k^2}.$$

В этой формуле k и $\varkappa$, как и $\boldsymbol{g}$, являются функциями параметра s.

Мы нашли короткую формулу для вычисления кручения кривой, но ею можно пользоваться, только если кривая снабжена естественной параметризацией, а это бывает на практике крайне редко. Выведем формулу для кручения при произвольной параметризации.

Пусть $\boldsymbol{f}(t)$ — произвольная, а $\boldsymbol{g}(s)$ — естественная параметризации кривой C, связанные заменой параметра $s = \psi(t) = \int\limits_a |\boldsymbol{f}'(\tau)|d\tau$. Тогда непосредственным вычислением убеждаемся в том, что

$$\begin{aligned}
\boldsymbol{f}'(t) &= \boldsymbol{g}'(s)\cdot\psi'(t),\\
\boldsymbol{f}''(t) &= \boldsymbol{g}''(s)\cdot[\psi'(t)]^2 + \boldsymbol{g}'(s)\psi''(t),\\
\boldsymbol{f}'''(t) &= \boldsymbol{g}'''(s)\cdot[\psi'(t)]^3 + 3\boldsymbol{g}''(s)\psi'(t)\psi''(t) + \boldsymbol{g}'(s)\psi'''(t).
\end{aligned}$$

Отсюда следует, что

$$(\boldsymbol{f}',\ \boldsymbol{f}'',\ \boldsymbol{f}''') = (\boldsymbol{g}',\ \boldsymbol{g}'',\ \boldsymbol{g}''')\cdot\psi'^6,$$

причем $\psi' = |\boldsymbol{f}'|$. Пользуясь этим и тем, что $k = |\boldsymbol{f}'\times\boldsymbol{f}''|/|\boldsymbol{f}'|^3$, из формулы для кручения в естественной параметризации получаем окончательно

$$\varkappa = \frac{(\boldsymbol{f}',\ \boldsymbol{f}'',\ \boldsymbol{f}''')}{|\boldsymbol{f}'\times\boldsymbol{f}''|^2}.$$

Это формула для вычисления кручения в произвольной параметризации.

§ 8. Натуральные уравнения кривой

С каждой гладкой кривой, кривизна которой во всех точках отлична от нуля, мы связали три функции: $s(t)$, $k(t)$, $\varkappa(t)$, которые могут быть найдены по произвольной параметризации $\boldsymbol{f}(t)$ этой кривой.

Спрашивается: определяют ли эти функции полностью кривую и нет ли между ними каких-либо соотношений?

Мы можем наложить на них только тривиальные ограничения. Функции эти непрерывны, кроме того, $k(t) > 0$, $s'(t) > 0$ при любом t. Легко видеть, что от трех функций можно перейти к двум, взяв за параметр длину дуги s. Тогда $k(s)$ и $\varkappa(s)$ будут функциями, связанными непосредственно с кривой (и направлением на ней). Нам остается узнать, определяют ли они кривую однозначно и нет ли между ними каких-либо соотношений (таких, как $\sin' = \cos$ и $\sin^2 + \cos^2 = 1$).

Теорема. *Пусть C_1 и C_2 — две гладкие кривые, имеющие одинаковую длину, и $\boldsymbol{g}_1(s)$, $\boldsymbol{g}_2(s)$ — их естественные параметризации. Если в соответствующих точках эти кривые имеют одинаковые кривизну и кручение:*

$$k_1(s) \equiv k_2(s), \quad \varkappa_1(s) \equiv \varkappa_2(s),$$

то существует наложение, переводящее кривую C_1 в кривую C_2. Другими словами, кривизна и кручение определяют кривую с точностью до положения в пространстве.

Доказательство. Совместим наложением реперы Френе обеих кривых в начальных точках, соответствующих значению параметра $s = 0$: пусть $\boldsymbol{g}_1(0) = \boldsymbol{g}_2(0)$, $\boldsymbol{t}_1(0) = \boldsymbol{t}_2(0)$, $\boldsymbol{n}_1(0) = \boldsymbol{n}_2(0)$, $\boldsymbol{b}_1(0) = \boldsymbol{b}_2(0)$. В этих условиях требуется доказать, что кривые совпадают: $\boldsymbol{g}_1(s) \equiv \boldsymbol{g}_2(s)$. Мы будем рассматривать $\boldsymbol{t}_1$, $\boldsymbol{t}_2$, $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, k_1, k_2, $\varkappa_1$, $\varkappa_2$ как функции параметра s, но для краткости всюду опустим аргумент (рис. 38).

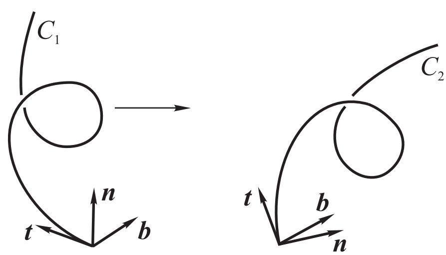


Рис. 38

Рассмотрим функцию $\xi(s)$, определенную по формуле

$$\xi(s) = \boldsymbol{t}_1 \cdot \boldsymbol{t}_2 + \boldsymbol{n}_1 \cdot \boldsymbol{n}_2 + \boldsymbol{b}_1 \cdot \boldsymbol{b}_2.$$

Докажем, что эта функция — постоянная. Для этого продифференцируем $\xi(s)$, пользуясь формулами Френе, и приведем подобные:

$$\xi'(s) = \boldsymbol{t}_1' \cdot \boldsymbol{t}_2 + \boldsymbol{t}_1 \cdot \boldsymbol{t}_2' + \boldsymbol{n}_1' \cdot \boldsymbol{n}_2 + \boldsymbol{n}_1 \cdot \boldsymbol{n}_2' + \boldsymbol{b}_1' \cdot \boldsymbol{b}_2 + \boldsymbol{b}_1 \cdot \boldsymbol{b}_2' =$$
$$= k_1 \boldsymbol{n}_1 \boldsymbol{t}_2 + k_2 \boldsymbol{t}_1 \boldsymbol{n}_2 - (k_1 \boldsymbol{t}_1 - \varkappa_1 \boldsymbol{b}_1)\boldsymbol{n}_2 -$$
$$(k_2 \boldsymbol{t}_2 - \varkappa_2 \boldsymbol{b}_2)\boldsymbol{n}_1 - \varkappa_1 \boldsymbol{n}_1 \boldsymbol{b}_2 - \varkappa_2 \boldsymbol{b}_1 \boldsymbol{n}_2 = 0.$$

Следовательно, $\xi(s) \equiv \xi(0) = 3$. Это означает, что $\boldsymbol{t}_1(s) \equiv \boldsymbol{t}_2(s)$, т. е. $\boldsymbol{g}\,'_1(s) \equiv \boldsymbol{g}\,'_2(s)$. Поскольку у вектор-функций $\boldsymbol{g}_1(s)$ и $\boldsymbol{g}_2(s)$ в каждой точке равны производные и совпадают их начальные значения ($\boldsymbol{g}_1(0) = \boldsymbol{g}_2(0)$), то эти функции равны: $\boldsymbol{g}_1(s) \equiv \boldsymbol{g}_2(s)$. Теорема доказана. □

Можно показать, что между кривизной и кручением нет нетривиальных соотношений. Более точно: *для любых гладких непрерывных функций* $h(s)$ *и* $\eta(s)$, *на отрезке* $[0,\ S]$, *первая из которых положительна, существует гладкая кривая* C, *кривизна и кручение которой в естественной параметризации определяются этими функциями:*

$$k = h(s), \quad \varkappa = \eta(s).$$

Доказательство этого факта мы не приводим. Оно использует теорему существования решения у системы обыкновенных дифференциальных уравнений.

Соотношения $k = h(s)$, $\varkappa = \eta(s)$ называются *натуральными уравнениями* кривой C. Их достоинство заключается в том, что они никак не зависят от выбора координат.

Как мы уже знаем, кривая C ими определена однозначно с точностью до положения в пространстве.

Глава II

ДИФФЕРЕНЦИАЛЬНАЯ ГЕОМЕТРИЯ ПОВЕРХНОСТЕЙ

§ 1. Элементарные поверхности в евклидовом пространстве. Способы их задания

Пусть Φ — множество в евклидовом пространстве $\mathbf{R}^3$. Множество Φ называется *элементарной поверхностью*, если при проекции на некоторую плоскость оно взаимно однозначно и взаимно непрерывно отображается на открытую область в этой плоскости (рис. 39).

Примеры. 1. Любая плоская область является элементарной поверхностью.

2. Сфера не является элементарной поверхностью, хотя таковой будет всякая достаточно малая сферическая область.

Если мы фиксируем плоскость, фигурирующую в определении элементарной поверхности, то получим возможность *явно задать* элементарную поверхность Φ. Пусть Π — рассматриваемая плоскость. Введем

в пространстве декартову систему координат xyz так, чтобы плоскость Π совпадала с координатной плоскостью xy. Тогда проекция на плоскость Π точки с координатами x, y, z будет иметь координаты x, y, 0. Если область $U \subset \Pi$ — образ элементарной поверхности Φ при проекции на Π, то множество Φ будет графиком некоторой непрерывной функции $f(x,\ y)$, определенной в области U. Поэтому множество Φ можно задать уравнением

$$z = f(x,\ y).$$

Такое задание называется *явным*, а само уравнение называется *уравнением элементарной поверхности* Φ *в явном виде* (рис. 40).

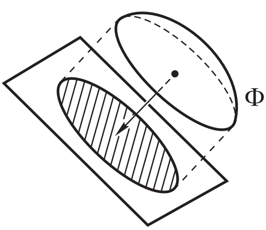


Рис. 39

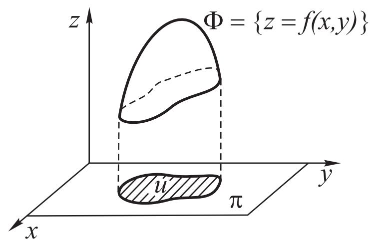


Рис. 40

Для многих целей элементарную поверхность Φ удобно рассматривать как образ области U при ее отображении в пространство $\mathbf{R}^3$, которое точке с координатами x, y ставит в соответствие точку с координатами x, y, $f(x,\ y)$. Это пример *параметрического задания* элементарной поверхности Φ.

Вообще, если элементарная поверхность Φ является образом плоской области V при непрерывном взаимно однозначном отображении F: $V \to \mathbf{R}^3$, то это отображение F называется *параметризацией* элементарной поверхности Φ. Элементарная поверхность, снабженная параметризацией, называется *параметризованной элементарной поверхностью.* Координаты в плоскости $\mathbf{R}^2$, в которой лежит область V, будем для определенности обозначать буквами u и v. Значениями двух параметров u и v полностью определяется положение любой точки P на поверхности: $P = F(u,\ v)$. Числа u и v будем также называть *внутренними координатами* точки P. Если точка P имеет координаты x, y, z, то при изменении параметров u и v координаты x, y, z тоже будут

меняться — каждая из них является некоторой функцией от u и v:

$$x = f_1(u,\ v), \quad y = f_2(u,\ v), \quad z = f_3(u,\ v), \tag{1}$$

где $f_1,\ f_2,\ f_3$ — непрерывные числовые функции, заданные в области V. Функции $f_1,\ f_2,\ f_3$ полностью описывают параметризацию F и называются ее *координатными функциями*. Соотношения $x = f_1$, $y = f_2$, $z = f_3$ называются *уравнениями* параметризованной элементарной поверхности Ф (рис. 41).

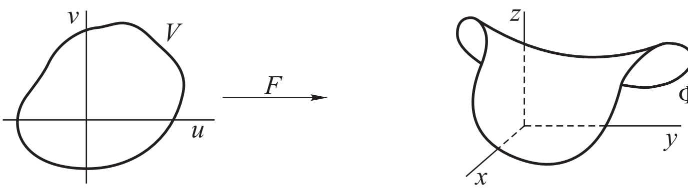


Рис. 41

Для нас основной интерес будут представлять элементарные поверхности, обладающие параметризацией с некоторыми дополнительными свойствами. Пусть $F\colon V \to \mathbf{R}^3$ — параметризация элементарной поверхности Ф, а $f_1,\ f_2,\ f_3$ — ее координатные функции. Если, во-первых, функции $f_1,\ f_2,\ f_3$ непрерывно дифференцируемы и, во-вторых, при каждом значении параметров u и v в точке $(u,\ v)$ не обращается в нуль по крайней мере один из трех определителей

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_2}{\partial u} \\ \dfrac{\partial f_1}{\partial v} & \dfrac{\partial f_2}{\partial v} \end{vmatrix}, \quad \begin{vmatrix} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_3}{\partial u} \\ \dfrac{\partial f_1}{\partial v} & \dfrac{\partial f_3}{\partial v} \end{vmatrix}, \quad \begin{vmatrix} \dfrac{\partial f_2}{\partial u} & \dfrac{\partial f_3}{\partial u} \\ \dfrac{\partial f_2}{\partial v} & \dfrac{\partial f_3}{\partial v} \end{vmatrix},$$

то параметризация F называется *регулярной*.

Элементарная поверхность, обладающая регулярной параметризацией, называется *гладкой*. При необходимости мы будем требовать от поверхности без дополнительных оговорок *повышенной гладкости*, т. е. существования у функций $f_1,\ f_2,\ f_3$ непрерывных частных производных всех порядков до n включительно, при некотором $n \geqslant 2$.

В случае явно заданной элементарной поверхности имеется простой достаточный критерий гладкости. Если $z = f(x,\ y)$ — явное уравнение элементарной поверхности Ф, то для того, чтобы Ф была гладкой, достаточно, чтобы функция f была непрерывно дифференцируемой. (Обратное, вообще говоря, неверно.)

Рассмотрим произвольное взаимно однозначное и взаимно непрерывное отображение φ некоторой плоской области W на область V (рис. 42):

$$\varphi: \; W \to V, \quad (\xi, \; \eta) \longmapsto (u, \; v),$$

где $u = \varphi_1(\xi, \; \eta)$, $v = \varphi_2(\xi, \; \eta)$. Если F: $V \to \mathbf{R}^3$ — параметризация элементарной поверхности Φ, то сквозное отображение G: $W \to \mathbf{R}^3$, определенное по формуле $G(\xi, \; \eta) = F(\varphi_1(\xi, \; \eta), \; \varphi_2(\xi, \; \eta))$, также будет параметризацией для Φ. Говорят, что она получается из параметризации F при помощи *замены внутренних координат* $u = \varphi_1(\xi, \; \eta)$, $v = \varphi_2(\xi, \; \eta)$.

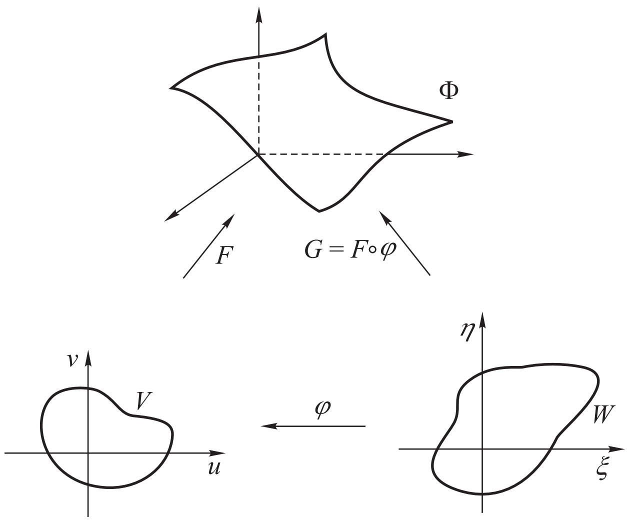


Рис. 42

Для того чтобы параметризация $G(\xi, \; \eta)$ элементарной поверхности Φ, полученная заменой $(u, \; v) = \varphi(\xi, \; \eta)$ из регулярной параметризации F, была также регулярной, необходимо и достаточно, чтобы замена была *неособой*, т. е. чтобы функции φ_1 и φ_2 были непрерывно дифференцируемы и якобиан замены не обращался в нуль:

$$\begin{vmatrix} \dfrac{\partial \varphi_1}{\partial u} & \dfrac{\partial \varphi_2}{\partial u} \\ \dfrac{\partial \varphi_1}{\partial v} & \dfrac{\partial \varphi_2}{\partial v} \end{vmatrix} \neq 0.$$

(Для доказательства воспользуйтесь правилом дифференцирования сложных функций.)

Всюду ниже в этой главе мы для краткости будем пользоваться термином ***поверхность***, имея при этом в виду только гладкие элементарные поверхности.

§ 2. Вектор-функции двух переменных

В предыдущем параграфе мы рассмотрели различные способы задания поверхностей. Они связаны с некоторой системой координат в пространстве и используют числовые функции двух переменных. Как и в случае кривых, часто мы будем пользоваться бескоординатным способом задания, при котором для параметризации поверхности служит вектор-функция двух переменных.

Вектор-функция двух переменных определена в некоторой области $W \subset \boldsymbol{R}^2$ и ставит в соответствие каждой точке $(x,\ y) \in W$ вектор $\boldsymbol{v}(x,\ y)$ трехмерного пространства (рис. 43). Все, что касается пределов, непрерывности и алгебраических операций над такими вектор-функциями, дословно повторяет уже известное нам о вектор-функциях одной переменной. Так же вводятся координатные функции v_1, v_2, v_3 (рис. 44). Небольшие отличия касаются дифференцирования: вместо одной производной у функции двух переменных есть две *частные* производные. Они обозначаются путем добавления к обозначению исходной функции нижних индексов, соответствующих переменным, по которым производится дифференцирование: $\boldsymbol{v}_x(x,\ y)$, $\boldsymbol{v}_y(x,\ y)$, а также $\boldsymbol{v}'_x(x,\ y)$, $\partial \boldsymbol{v}/\partial y$, $\partial_x \boldsymbol{v}$ и т. п. Отметим, что если вместо x и y, стоящих в скобках, можно подставлять конкретные числа, то x и y, стоящие в качестве индексов, образуют *единый символ* с обозначением функции.

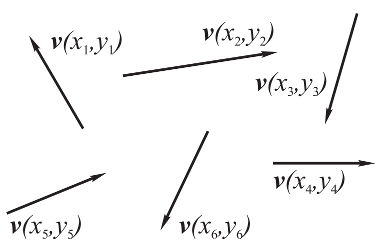


Рис. 43

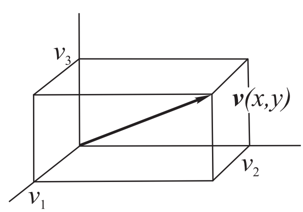


Рис. 44

Частные производные координатных функций функции $\boldsymbol{v}(x,\ y)$ совпадают с кооодинатными функциями ее частных производных.

Если $F\colon V \to \boldsymbol{R}^3$ — параметризация поверхности Φ, то вектор-

функция $\boldsymbol{f}$, определенная по формуле

$$\boldsymbol{f}(u, v) = \overrightarrow{OF(u, v)},$$

называется векторной параметризацией поверхности Φ (рис. 45), а соотношение

$$\boldsymbol{r} = \boldsymbol{f}(u, v)$$

называется ее *векторным уравнением.* В силу непрерывности функции F вектор-функция $\boldsymbol{f}$ также непрерывна. Если поверхность Φ гладкая, а F — ее регулярная параметризация, то функция $\boldsymbol{f}$ непрерывно дифференцируема в области V, причем ее частные производные в каждой точке линейно независимы:

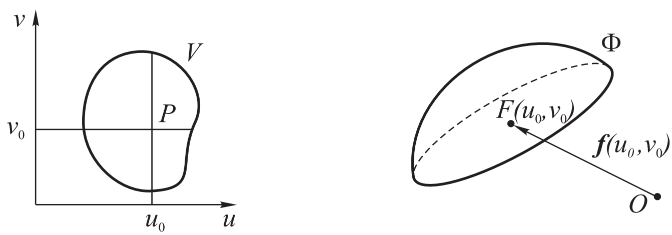


Рис. 45

$$\boldsymbol{f}_u(u, v) \times \boldsymbol{f}_v(u, v) \neq 0.$$

В дальнейшем мы будем пользоваться только векторными параметризациями, причем, как и в случае кривых, не будем различать точку и ее радиус-вектор и будем использовать запись $\boldsymbol{f}(u, v) = P$ вместо $\boldsymbol{f}(u, v) = \overrightarrow{OP}$.

§ 3. Кривые на гладкой поверхности

Пусть Φ — гладкая поверхность, заданная уравнением

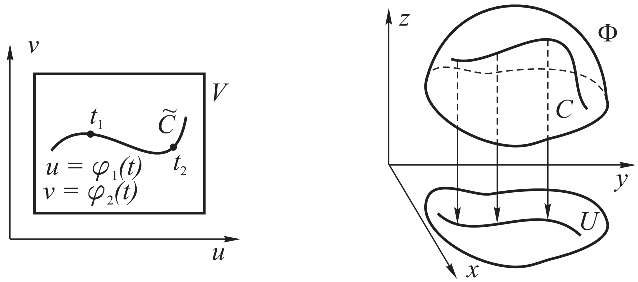
$$\boldsymbol{r} = \boldsymbol{f}(u, v),$$

где $\boldsymbol{f}\colon V \to \boldsymbol{R}^3$ — ее регулярная параметризация. Пусть $u = \varphi_1(t)$, $v = \varphi_2(t)$ — уравнения некоторой параметризованной кривой $\widetilde{C}$ в области V. На поверхности Φ кривой $\widetilde{C}$ соответствует кривая $C = \boldsymbol{f}(\widetilde{C})$,

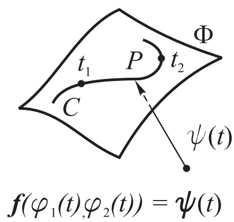
которая в пространстве задается векторным уравнением

$$\boldsymbol{r} = \boldsymbol{\psi}(t),$$

где $\boldsymbol{\psi}(t) = \boldsymbol{f}(\varphi_1(t),\ \varphi_2(t))$ (рис. 46).



а



б

Рис. 46

Теорема. *Если кривая $\widetilde{C}$ гладкая, то и кривая C тоже будет гладкой.*

Доказательство. Во-первых, ясно, что если функции $\boldsymbol{f}(u,\ v)$, $\varphi_1(t)$ и $\varphi_2(t)$ n раз непрерывно дифференцируемы, то и их композиция $\boldsymbol{\psi}(t)$ будет n раз непрерывно дифференцируемой. Далее, воспользуемся равенством

$$\boldsymbol{\psi}'(t) = \boldsymbol{f}_u(\varphi_1(t),\ \varphi_2(t)) \cdot \varphi_1'(t) + \boldsymbol{f}_v(\varphi_1(t),\ \varphi_2(t)) \cdot \varphi_2'(t). \qquad (2)$$

Так как в любой точке $(u,\ v)$ векторы частных производных $\boldsymbol{f}_u$ и $\boldsymbol{f}_v$ линейно независимы, и при любом t хотя бы одно из чисел $\varphi_1'(t)$, $\varphi_2'(t)$ не равно нулю, то и $\boldsymbol{\psi}'(t) \neq \boldsymbol{0}$. Следовательно, $\boldsymbol{\psi}'(t)$ есть регулярная

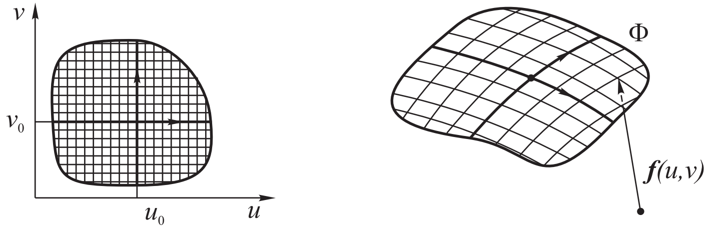


Рис. 47

параметризация кривой C, которая, таким образом, является гладкой. Теорема доказана. □

Параметрические уравнения кривой $\widetilde{C}$ в области V

$$u = \varphi_1(t), \quad v = \varphi_2(t)$$

называются *внутренними уравнениями* кривой C на поверхности Φ.

Особый интерес представляют кривые, которые являются образами отрезков в области V, параллельных осям координат (рис. 47). Они задаются внутренними уравнениями вида

$$u = t,\ v = v_0 = \text{const};$$
$$u = u_0 = \text{const},\ v = t$$

и называются *координатными линиями* на параметризованной поверхности Φ. Мы теперь можем дать геометрическое истолкование векторов частных производных функции $\boldsymbol{f}(u,\ v)$. Из формулы (2) видно, что векторы $\boldsymbol{f}_u(u_0,\ v_0)$ и $\boldsymbol{f}_v(u_0,\ v_0)$ — касательные к координатным линиям в точке $P = \boldsymbol{f}(u_0,\ v_0)$. Поскольку поверхность Φ гладкая, а $\boldsymbol{f}(u,\ v)$ — ее регулярная параметризация, то векторы $\boldsymbol{f}_u(u_0,\ v_0)$ и $\boldsymbol{f}_v(u_0,\ v_0)$ неколлинеарны, и угол между ними равен углу между координатными линиями в точке P (рис. 48).

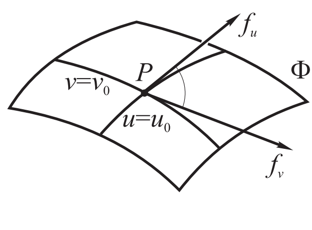


Рис. 48

В дальнейшем мы будем рассматривать только гладкие кривые на поверхности, заданные своими регулярными параметризациями.

§ 4. Касательная плоскость поверхности

Пусть $\Phi \subset \boldsymbol{R}^3$ — гладкая поверхность, а P — некоторая ее точка. Говорят, что прямая *касается* поверхности Φ в точке P, если она является касательной прямой в точке P некоторой кривой, лежащей в поверхности Φ и проходящей через точку P (рис. 49).

Теорема 1. *Все прямые, касающиеся поверхности* Φ *в точке* P, *лежат в одной плоскости.*

Доказательство. Пусть поверхность Φ задана уравнением

$$\boldsymbol{r} = \boldsymbol{f}(u,\ v).$$

Тогда, как видно из формулы (2) § 3, касательный вектор в точке $P = \boldsymbol{f}(u_0,\ v_0)$ любой кривой, проходящей через точку P и лежащей в поверхности Φ, является линейной комбинацией векторов частных

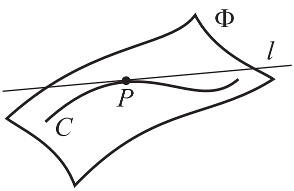


Рис. 49

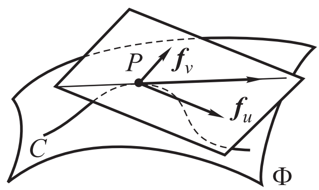


Рис. 50

производных $\boldsymbol{f}_u(u_0,\ v_0)$ и $\boldsymbol{f}_v(u_0,\ v_0)$. Следовательно, каждый такой вектор, если его отложить из точки P, будет лежать в плоскости, проходящей через эту точку и содержащей векторы $\boldsymbol{f}_u(u_0,\ v_0)$ и $\boldsymbol{f}_v(u_0,\ v_0)$ (рис. 50). Ясно, что в этой же плоскости лежат и все рассматриваемые касательные прямые. Теорема доказана. □

Плоскость, в которой лежат все прямые, касающиеся поверхности Φ в точке P, называется *касательной плоскостью*. Мы будем обозначать ее через $T_P\Phi$.

Если $\boldsymbol{r} = \boldsymbol{f}(u,\ v)$ — уравнение поверхности, то в качестве нормального вектора к касательной плоскости в точке $P = \boldsymbol{f}(u_0,\ v_0)$ естественно взять векторное произведение $\boldsymbol{f}_u(u_0,\ v_0) \times \boldsymbol{f}_v(u_0,\ v_0)$, поскольку векторы $\boldsymbol{f}_u(u_0,\ v_0)$ и $\boldsymbol{f}_v(u_0,\ v_0)$ заведомо в ней лежат (рис. 51). В этом случае векторное уравнение касательной плоскости имеет вид

$$(\boldsymbol{r} - \boldsymbol{f}(u_0,\ v_0),\ \boldsymbol{f}_u(u_0,\ v_0),\ \boldsymbol{f}_v(u_0,\ v_0)) = 0.$$

В координатах это уравнение принимает вид

$$\begin{vmatrix} X - x_0 & Y - y_0 & Z - z_0 \\ \partial_u f_1(u_0,\ v_0) & \partial_u f_2(u_0,\ v_0) & \partial_u f_3(u_0,\ v_0) \\ \partial_v f_1(u_0,\ v_0) & \partial_v f_2(u_0,\ v_0) & \partial_v f_3(u_0,\ v_0) \end{vmatrix} = 0,$$

где $x_0 = f_1(u_0,\ v_0)$, $y_0 = f_2(u_0,\ v_0)$, $z_0 = f_3(u_0,\ v_0)$. *Единичный вектор нормали* к поверхности в точке P определяется по формуле

$$\boldsymbol{n} = \frac{\boldsymbol{f}_u \times \boldsymbol{f}_v}{|\boldsymbol{f}_u \times \boldsymbol{f}_v|}.$$

Эта формула задает нормаль поверхности Φ как вектор-функцию внутренних координат u и v (рис. 52).

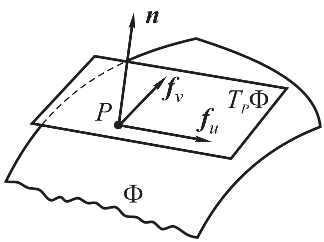


Рис. 51

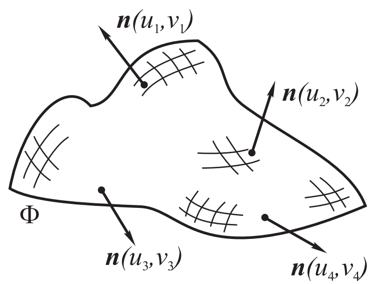


Рис. 52

Касательная плоскость с точностью до величин первого порядка малости приближает поверхность в данной ее точке. Более точно, пусть Q — точка поверхности Φ, близкая к точке P. *При стремлении точки Q к точке P отношение расстояния h от точки Q до касательной плоскости $T_P\Phi$ к расстоянию от Q до P стремится к нулю:*

$$\lim_{Q \to P} \frac{h}{QP} = 0.$$

Касательная плоскость — единственная, обладающая этим свойством (рис. 53).

Доказывать этого мы не будем.

С касательной плоскостью в точке P связано одно очень удобное явное задание поверхности Φ в некоторой малой окрестности этой точки.

Теорема 2. *Пусть* Φ — *гладкая поверхность* и P — *произвольная ее точка. Рассмотрим декартову систему координат, начало которой совпадает с точкой* P, *а ось* z *направлена по нормали к поверхности.* (*При этом оси* x *и* y *окажутся лежащими в касательной плоскости* $T_P\Phi$.) *Тогда у точки* P *найдется окрестность в поверхности, которую в координатах* x, y, z *можно задать явным уравнением*

$$z = f(x,\ y),$$

где функция $f(x,\ y)$ *определена и непрерывно дифференцируема в некоторой достаточно малой окрестности точки* $(0, 0)$ *на плоскости* x, y, *причем в самой точке* $(0,\ 0)$ *имеют место соотношения*

$$f(0,\ 0) = f_x(0,\ 0) = f_y(0,\ 0) = 0.$$

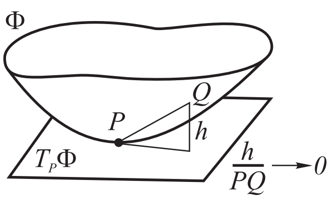


Рис. 53

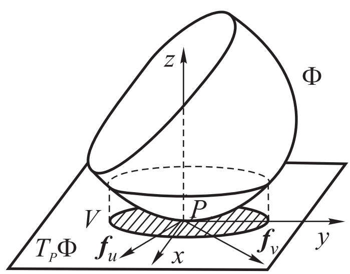


Рис. 54

Кроме того, если поверхность Φ *допускает* n *раз непрерывно дифференцируемую регулярную параметризацию, то функция* $f(x,\ y)$ *тоже* n *раз непрерывно дифференцируема* (рис. 54).

Д о к а з а т е л ь с т в о. Пусть поверхность Φ задана уравнениями

$$\begin{cases} x = f_1(u,\ v), \\ y = f_2(u,\ v), \\ z = f_3(u,\ v), \end{cases}$$

причем $f_1(u_0,\ v_0) = f_2(u_0,\ v_0) = f_3(u_0,\ v_0) = 0$. Рассмотрим проекцию поверхности Φ на плоскость xy. При этой проекции точка с внутренними координатами $(u,\ v)$ отображается в точку с координатами $f_1(u,\ v)$,

$f_2(u,\ v)$. Рассмотрим векторы частных производных (где $\partial_u f_1 = (f_1)_u$ и т. д.):

$$\boldsymbol{f}_u = (\partial_u f_1,\ \partial_u f_2,\ \partial_u f_3),$$
$$\boldsymbol{f}_v = (\partial_v f_1,\ \partial_v f_2,\ \partial_v f_3).$$

Поскольку, в силу выбора системы координат, векторы $\boldsymbol{f}_u(u_0,\ v_0)$, $\boldsymbol{f}_v(u_0,\ v_0)$ лежат в плоскости xy, то

$$\partial_u f_3(u_0,\ v_0) = \partial_v f_3(u_0,\ v_0) = 0.$$

Следовательно, в силу линейной независимости векторов $\boldsymbol{f}_u(u_0,\ v_0)$ и $\boldsymbol{f}_v(u_0,\ v_0)$ получаем

$$\begin{vmatrix} \partial_u f_1(u_0,\ v_0) & \partial_u f_2(u_0,\ v_0) \\ \partial_v f_1(u_0,\ v_0) & \partial_v f_2(u_0,\ v_0) \end{vmatrix} \neq 0.$$

Если мы рассмотрим теперь отображение области V в плоскость xy, заданное формулами

$$x = f_1(u,\ v), \quad y = f_2(u,\ v),$$

то из теоремы об обратной функции будет следовать, что в некоторой окрестности U_0 точки $(0,\ 0)$ определено обратное отображение $u = \varphi_1(x,\ y)$, $v = \varphi_2(x,\ y)$. Теперь ясно, что искомую функцию f можно определить по формуле

$$f(x,\ y) = f_3(\varphi_1(x,\ y),\ \varphi_2(x,\ y)).$$

То, что $f(0,\ 0) = 0$, очевидно. То, что $f_x(0,\ 0) = f_y(0,\ 0) = 0$, следует из того, что в точке $(u_0,\ v_0) = (\varphi_1(0,\ 0),\ \varphi_2(0,\ 0))$ частные производные функции f_3 обращаются в нуль. Наконец, если параметризация $f(u,\ v)$ непрерывно дифференцируема n раз, то вместе с ее координатными функциями $f_1,\ f_2$, n раз непрерывно дифференцируемы функции φ_1 и φ_2. Теперь функция f непрерывно дифференцируема n раз как композиция функций φ_1, φ_2 и n раз непрерывно дифференцируемой функции f_3. Теорема 2 доказана. □

Направления на поверхности. Пусть l — некоторая прямая, проходящая через точку P поверхности и лежащая в касательной плоскости $T_P\Phi$. Будем говорить, что кривая C а поверхности проходит через точку P *в направлении* прямой l, если l — касательная кривой C в точке P (рис. 49). Если через точку P проходят две кривые C_1 и

C_2, то *углом между ними* назовем угол между их касательными в этой точке. Очевидно, что угол между кривыми зависит только от их направлений в точке P (рис. 55).

Рассмотрим плоскость, перпендикулярную касательной плоскости $T_P\Phi$ и пересекающую ее по прямой l, которая по-прежнему проходит через P. Из доказанной выше теоремы 2 следует, что пересечение этой

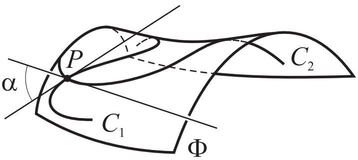


Рис. 55

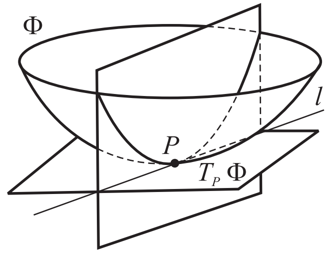


Рис. 56

плоскости с достаточно малой окрестностью точки P на поверхности является гладкой кривой (это график сужения функции f из теоремы на прямую l). Эта кривая называется *нормальным сечением* поверхности Φ в точке P в направлении прямой l. Очевидно, что l — касательная прямая этого нормального сечения. Более общим образом, касательная к кривой на поверхности всегда лежит в пересечении соприкасающейся плоскости кривой и касательной плоскости поверхности (рис. 56). (Проверьте это!)

§ 5. Первая квадратичная форма поверхности. Измерение длин кривых и углов между ними

Длина кривой на поверхности. Пусть Φ — гладкая поверхность, заданная векторным уравнением $\boldsymbol{r} = \boldsymbol{f}(u,\ v)$. Рассмотрим гладкую кривую C на поверхности Φ, заданную внутренними уравнениями $u = \varphi_1(t)$, $v = \varphi_2(t)$, где $t \in [a,\ b]$. Нас интересует длина S кривой C. В пространстве кривая C задана векторным уравнением $\boldsymbol{r} = \boldsymbol{\varphi}(t)$, где $\boldsymbol{\varphi}(t) = \boldsymbol{f}(\varphi_1(t),\ \varphi_2(t))$. Теперь длину S можно найти по известной формуле, см. § I. 4:

$$S = \int\limits_a^b |\boldsymbol{\varphi}'(t)|dt.$$

Вычислим длину вектора $\boldsymbol{\varphi}'(t)$. По формуле (2), § 3

$$\boldsymbol{\varphi}'(t) = \boldsymbol{f}_u(\varphi_1(t),\ \varphi_2(t)) \cdot \varphi_1'(t) + \boldsymbol{f}_v(\varphi_1(t),\ \varphi_2(t)) \cdot \varphi_2'(t),$$

или, более коротко,

$$\boldsymbol{\varphi}' = \boldsymbol{f}_u(\varphi_1,\ \varphi_2) \cdot \varphi_1' + \boldsymbol{f}_v(\varphi_1,\ \varphi_2) \cdot \varphi_2'.$$

Далее получаем, для краткости опуская t,

$$\begin{aligned} &|\boldsymbol{\varphi}'(t)|^2 = |\boldsymbol{\varphi}'|^2 = \boldsymbol{\varphi}' \cdot \boldsymbol{\varphi}' = \\ &\qquad = \boldsymbol{f}_u^2(\varphi_1, \varphi_2)\varphi_1'^2 + 2\boldsymbol{f}_u(\varphi_1, \varphi_2)\boldsymbol{f}_v(\varphi_1, \varphi_2)\varphi_1'\varphi_2' + \boldsymbol{f}_v^2(\varphi_1, \varphi_2)\varphi_2'^2. \end{aligned}$$

Введем обозначения. Положим

$$\begin{aligned} E(u,\ v) &= \boldsymbol{f}_u^2(u,\ v), \\ F(u,\ v) &= \boldsymbol{f}_u(u,\ v) \cdot \boldsymbol{f}_v(u,\ v), \\ G(u,\ v) &= \boldsymbol{f}_v^2(u,\ v). \end{aligned}$$

Тогда формула для квадрата длины вектора $\boldsymbol{\varphi}'(t)$ примет вид

$$|\boldsymbol{\varphi}'(t)|^2 = E(\varphi_1,\ \varphi_2)\varphi_1'^2 + 2F(\varphi_1,\ \varphi_2)\varphi_1'\varphi_2' + G(\varphi_1,\ \varphi_2)\varphi_2'^2.$$

Наконец, подставляя корень из этого выражения в формулу длины, получим:

$$S = \int\limits_a^b \sqrt{E(\varphi_1,\ \varphi_2)\varphi_1'^2 + 2F(\varphi_1,\ \varphi_2)\varphi_1'\varphi_2' + G(\varphi_1,\ \varphi_2)\varphi_2'^2}\,dt. \qquad (*)$$

Имея в виду внутренние уравнения кривой C и опуская для краткости аргументы у функций E, F, G, эту формулу часто записывают так:

$$\begin{aligned} S = \int\limits_a^b \sqrt{E\left(\frac{du}{dt}\right)^2 + 2F\frac{du}{dt}\frac{dv}{dt} + G\left(\frac{dv}{dt}\right)^2}\,dt = \\ = \int\limits_a^b \sqrt{Eu_t^2 + 2Fu_tv_t + Gv_t^2}\,dt. \end{aligned}$$

(Но нельзя забывать, что E, F и G — это не числа, а *числовые функции*, зависящие от внутренних координат u и v!)

Первая квадратичная форма поверхности. На первый взгляд, формула (∗) кажется сложнее, чем исходное выражение $S = \int\limits_a^b |\boldsymbol{\varphi}'|dt$. Однако полученная формула (∗) имеет следующее важное преимущество: в ней выделены функции E, F и G, которые не зависят от выбора кривой C, а зависят только от поверхности и ее параметризации. Если ищутся длины сразу нескольких различных кривых на поверхности, то удобно сначала найти функции E, F и G, а потом подставлять в формулу (∗) функции $\varphi_1(t)$, $\varphi_2(t)$ и пределы интегрирования, которые отвечают той или другой кривой.

Кроме того, эта формула дает возможность говорить о длине кривой на поверхности, когда известны ее внутренние уравнения и функции E, F и G, а точный вид параметризации неизвестен (т. е. не задан точный вид поверхности). Это важно при общем изучении поверхностей (ср. § 12).

Определение. Функции E, F, G называются *коэффициентами первой квадратичной формы* поверхности Φ, а сама *первая квадратичная форма* I определяется по формуле

$$\mathrm{I}(\varphi_1', \ \varphi_2') = E \cdot \varphi_1'^2 + 2F \cdot \varphi_1'\varphi_2' + G \cdot \varphi_2'^2.$$

(Другое название — первая основная или первая фундаментальная форма поверхности.) Можно заметить, что матрица

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}$$

является матрицей скалярного произведения в касательной плоскости к поверхности, записанной в базисе $\boldsymbol{f}_u$, $\boldsymbol{f}_v$.

Пример. Для поверхности, заданной явным уравнением $z = f(x, \ y)$, коэффициенты первой квадратичной формы имеют вид:

$$\begin{aligned} E(x, \ y) &= 1 + f_x^2(x, \ y), \\ F(x, \ y) &= f_x(x, \ y) \cdot f_y(x, \ y), \\ G(x, \ y) &= 1 + f_y^2(x, \ y). \end{aligned}$$

Угол между кривыми на поверхности. Пусть C_1 и C_2 — две кривые на поверхности Φ, заданные своими внутренними уравнениями

$$u = \varphi_1(t), \ v = \varphi_2(t); \quad u = \psi_1(t), \ v = \psi_2(t).$$

Если C_1 и C_2 проходят через одну и ту же точку $P = \boldsymbol{f}(u_0,\ v_0)$ на поверхности, то при помощи коэффициентов E, F, G первой квадратичной формы можно найти угол θ между этими кривыми в точке P (он равен углу между касательными прямыми этих кривых). Если $u_0 = \varphi_1(t_0) = \psi_1(\tau_0)$, $v_0 = \varphi_2(t_0) = \psi_2(\tau_0)$, то

$$\cos\theta = \frac{\boldsymbol{\varphi}'(t_0) \cdot \boldsymbol{\psi}'(\tau_0)}{|\boldsymbol{\varphi}'(t_0)| \cdot |\boldsymbol{\psi}'(\tau_0)|},$$

где, как и раньше, $\boldsymbol{\varphi}' = \boldsymbol{f}_u \cdot \varphi_1' + \boldsymbol{f}_v \cdot \varphi_2'$ и $\boldsymbol{\psi}' = \boldsymbol{f}_u \psi_1' + \boldsymbol{f}_v \psi_2'$ — касательные векторы к кривым C_1 и C_2. Положим для краткости

$$\boldsymbol{\varphi}' = \boldsymbol{\varphi}'(t_0),\ \varphi_1' = \varphi_1'(t_0),\ \varphi_2' = \varphi_2'(t_0);$$
$$\boldsymbol{\psi}' = \boldsymbol{\psi}'(\tau_0),\ \psi_1' = \psi_1'(\tau_0),\ \psi_2' = \psi_2'(\tau_0);$$
$$\boldsymbol{f}_u = \boldsymbol{f}_u(u_0,\ v_0),\ \boldsymbol{f}_v = \boldsymbol{f}_v(u_0,\ v_0);$$
$$E = E(u_0,\ v_0),\ F = F(u_0,\ v_0),\ G = G(u_0,\ v_0).$$

Тогда скалярное произведение в числителе рассматриваемой дроби можно найти по формуле

$$\boldsymbol{\varphi}' \cdot \boldsymbol{\psi}' = (\varphi_1' \boldsymbol{f}_u + \varphi_2' \boldsymbol{f}_v) \cdot (\psi_1' \boldsymbol{f}_u + \psi_2' \boldsymbol{f}_v) = \\ = E \cdot \varphi_1' \psi_1' + F(\varphi_1' \psi_2' + \varphi_2' \psi_1') + G \varphi_2' \psi_2'.$$

Длины векторов $\boldsymbol{\varphi}'$ и $\boldsymbol{\psi}'$ находятся по формулам

$$|\boldsymbol{\varphi}'| = \sqrt{E\varphi_1'^2 + 2F\varphi_1'\varphi_2' + G\varphi_2'^2} = \sqrt{\mathrm{I}(\varphi_1',\ \varphi_2')},$$
$$|\boldsymbol{\psi}'| = \sqrt{E\psi_1'^2 + 2F\psi_1'\psi_2' + G\psi_2'^2} = \sqrt{\mathrm{I}(\psi_1',\ \psi_2')}.$$

Выясним теперь геометрический смысл коэффициентов первой квадратичной формы. E и G представляют собой квадраты масштабов координатных линий — длины малых дуг координатных u- и v-линий примерно в $\sqrt{E}$ и $\sqrt{G}$ раз больше соответствующих дуг в области V. Геометрический смысл коэффициента F сложнее:

$$F = \boldsymbol{f}_u \cdot \boldsymbol{f}_v = |\boldsymbol{f}_u| \cdot |\boldsymbol{f}_v| \cdot \cos\omega,$$

где ω — угол между координатными линиями. Отсюда

$$\cos\omega = \frac{F}{\sqrt{EG}}.$$

В частности, чтобы координатные линии в каждой точке были ортогональны, необходимо и достаточно, чтобы $F \equiv 0$.

Выясним геометрический смысл выражения $\sqrt{EG - F^2}$, которое понадобится нам в дальнейшем:

$$EG - F^2 = \boldsymbol{f}_u^2 \boldsymbol{f}_v^2 - \boldsymbol{f}_u^2 \boldsymbol{f}_v^2 \cos^2 \omega = \boldsymbol{f}_u^2 \boldsymbol{f}_v^2 \sin^2 \omega = (\boldsymbol{f}_u \times \boldsymbol{f}_v)^2.$$

Таким образом, число $\sqrt{EG - F^2}$ равняется площади параллелограмма, натянутого на векторы частных производных $\boldsymbol{f}_u$ и $\boldsymbol{f}_v$ (рис. 57).

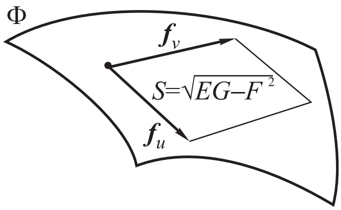


Рис. 57

Внутренняя геометрия поверхности. Подведем итоги. Как мы видели, для определения длины кривой на поверхности совершенно не нужно знать, что собой представляет поверхность. Достаточно знать ее первую квадратичную форму. Это же относится и к углу между кривыми. Как мы увидим в § 10, площади фигур на поверхности тоже могут быть найдены только по коэффициентам E, F и G. Говорят, что первая квадратичная форма отвечает за *внутреннюю геометрию* поверхности. (Подробнее см. § 12.)

На поверхности внутренние координаты можно выбирать многими способами. С точки зрения вычислений было бы разумно выбрать их так, чтобы упростить коэффициенты первой квадратичной формы.

Оказывается, что на всякой поверхности координаты можно ввести так, чтобы $E(u,\ v) \equiv 1$, $F(u,\ v) \equiv 0$. При этом функция $G(u,\ v)$, в принципе, может быть любой положительной. Такая система называется *полугеодезической* (см. § 15.)

Другой удобный вид координат — *изотермические*, или *конформные*, координаты. В этой системе $F(u,\ v) \equiv 0$, $E(u,\ v) \equiv G(u,\ v)$. Эти координаты замечательны тем, что лежащие в области V прообразы «маленьких фигур» на поверхности Φ почти подобны им, а углы между кривыми сохраняются.

Наконец, упомянем *чебышевские* координаты: в них $E(u,\ v) \equiv G(u,\ v) \equiv 1$. Здесь длины координатных линий вообще не меняются, а $F(u,\ v)$ равняется косинусу угла между координатными линиями: $F = \cos\omega$. Эти координаты применяются в задачах, связанных с раскроем (ткани, металла и т. п.).

§ 6. Кривизна кривой на поверхности. Вторая квадратичная форма

Пусть Φ — гладкая поверхность, параметризованная вектор-функцией $\boldsymbol{f}(u,\ v)$, а C — гладкая кривая на этой поверхности, заданная своими внутренними уравнениями $u = \varphi_1(t)$, $v = \varphi_2(t)$, которым соответствует ее параметризация $\boldsymbol{\varphi}(t) = \boldsymbol{f}(\varphi_1,\ \varphi_2)$. Пусть $P = \boldsymbol{\varphi}(t_0)$ — точка на кривой C. Найдем формулу для кривизны кривой C в точке P. Обозначим через θ угол между нормалью $\boldsymbol{n}$ поверхности Φ в точке P и главной нормалью $\boldsymbol{m}$ кривой C в точке P (рис. 58).

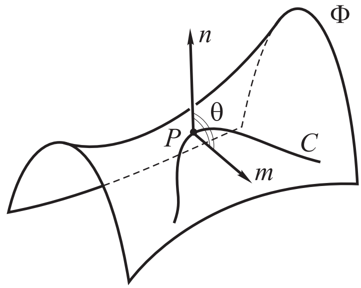


Рис. 58

Рассмотрим естественную параметризацию $\boldsymbol{\psi}(s)$ кривой C. Пусть она соответствует внутренним уравнениям $u = \psi_1(s)$, $v = \psi_2(s)$, и пусть при этом $P = \boldsymbol{\psi}(s_0)$. Кривизна k кривой C в точке P вычисляется по формуле (см. гл. I, § 5)

$$k = |\boldsymbol{\psi}''(s_0)| = \frac{\boldsymbol{\psi}''(s_0) \cdot \boldsymbol{n}}{\cos\theta}.$$

Чтобы найти $\boldsymbol{\psi}''(s_0)$, продифференцируем выражение для $\boldsymbol{\psi}'(s)$:

$$\boldsymbol{\psi}'(s) = \boldsymbol{f}_u(\psi_1,\ \psi_2)\psi_1' + \boldsymbol{f}_v(\psi_1,\ \psi_2)\psi_2'.$$

Проделав несложные вычисления, получим

$$\boldsymbol{\psi}''(s) = \boldsymbol{f}_{uu}(\psi_1,\ \psi_2)\psi_1'^2 + 2\boldsymbol{f}_{uv}(\psi_1,\ \psi_2)\psi_1'\psi_2' + \\ + \boldsymbol{f}_{vv}(\psi_1,\ \psi_2)\psi_2'^2 + \boldsymbol{f}_u(\psi_1,\ \psi_2)\psi_1'' + \boldsymbol{f}_v(\psi_1,\ \psi_2)\psi_2''.$$

При скалярном умножении на вектор $\boldsymbol{n}$ последние два слагаемых дадут нуль, так как первые производные вектор-функции $\boldsymbol{f}$ ортогональны вектору $\boldsymbol{n}$.

Для скалярных произведений вторых производных вектор-функции $\boldsymbol{f}$ на вектор $\boldsymbol{n}$ приняты специальные обозначения:

$$L(u,\ v) = \boldsymbol{f}_{uu}(u,\ v) \cdot \boldsymbol{n}(u,\ v),$$
$$M(u,\ v) = \boldsymbol{f}_{uv}(u,\ v) \cdot \boldsymbol{n}(u,\ v),$$
$$N(u,\ v) = \boldsymbol{f}_{vv}(u,\ v) \cdot \boldsymbol{n}(u,\ v),$$

или, короче,

$$L = \boldsymbol{f}_{uu} \cdot \boldsymbol{n}, \quad M = \boldsymbol{f}_{uv} \cdot \boldsymbol{n}, \quad N = \boldsymbol{f}_{vv} \cdot \boldsymbol{n}.$$

Пользуясь этими обозначениями, можно записать

$$k \cdot \cos\theta = \boldsymbol{\psi}''(s_0) \cdot \boldsymbol{n}(\psi_1(s_0),\ \psi_2(s_0)) = \boldsymbol{\psi}'' \cdot \boldsymbol{n} =$$
$$= L(\psi_1,\ \psi_2)\psi_1'^2 + 2M(\psi_1,\ \psi_2) \cdot \psi_1'\psi_2' + N(\psi_1,\ \psi_2)\psi_2'^2,$$

где значения функций ψ_1, ψ_2, ψ_1', ψ_2' берутся в точке s_0. Более краткая запись:

$$k \cdot \cos\theta = L \cdot \psi_1'^2 + 2M \cdot \psi_1'\psi_2' + N \cdot \psi_2'^2.$$

(Но нельзя забывать, что L, M, N — *числовые функции* двух переменных u и v!)

Вернемся к исходной параметризации $\boldsymbol{\varphi}(t)$. Поскольку $\boldsymbol{\psi}'(s) = \boldsymbol{\varphi}'(t)/|\boldsymbol{\varphi}'(t)|$, то $\psi_1' = \varphi_1'/|\boldsymbol{\varphi}'|$, $\psi_2' = \varphi_2'/|\boldsymbol{\varphi}'|$, где все значения берутся по-прежнему в s_0 и t_0. Подставляя в предыдущую формулу, получим

$$k \cdot \cos\theta = \frac{L \cdot \varphi_1'^2 + 2M \cdot \varphi_1'\varphi_2' + N \cdot \varphi_2'^2}{E \cdot \varphi_1'^2 + 2F \cdot \varphi_1'\varphi_2' + G \cdot \varphi_2'^2}.$$

Функции L, M, N называются *коэффициентами второй квадратичной формы*, а сама вторая форма определяется по формуле

$$\mathrm{II}(\varphi_1',\ \varphi_2') = L \cdot \varphi_1'^2 + 2M \cdot \varphi_1'\varphi_2' + N \cdot \varphi_2'^2.$$

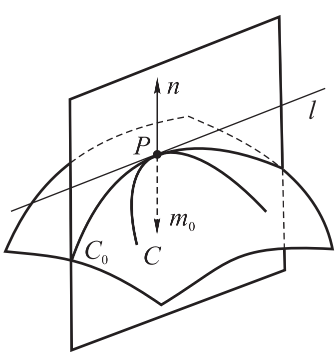


Рис. 59

Часто удобно считать вторую квадратичную форму функцией, определенной на множестве касательных векторов или просто на касательной плоскости $T_P\Phi$.

Теперь полученное выражение для кривизны кривой C в точке P принимает вид

$$k = \frac{\mathrm{II}(\varphi_1',\ \varphi_2')}{\mathrm{I}(\varphi_1',\ \varphi_2') \cdot \cos\theta}.$$

Нормальная кривизна поверхности. Особый интерес представляют кривые, для которых

$\cos\theta = \pm 1$, т. е. те кривые, чья соприкасающаяся плоскость в точке P перпендикулярна касательной плоскости $T_P\Phi$. Пусть l — касательная прямая кривой C в точке P. Рассмотрим плоскость, проходящую через l и нормаль к Φ в точке P. Ее пересечение с достаточно малой окрестностью точки P на поверхности Φ дает нам неособую кривую C_0 — нормальное сечение поверхности Φ в направлении касательной l. Если θ_0 — угол между главной нормалью $\boldsymbol{m}_0$ к C_0 и нормалью $\boldsymbol{n}$ к поверхности в точке P, то $\cos\theta_0 = \pm 1$. Пусть k_0 — кривизна этой кривой в точке P. Тогда число

$$k_n = k_0 \cdot \cos\theta_0 = \mathrm{II}/\mathrm{I}$$

называется *нормальной кривизной* поверхности Φ в направлении касательной l (рис. 59).

Теорема 1 (Мёнье). *Кривизна k кривой C на поверхности зависит только от угла θ и нормальной кривизны k_n в направлении касательной этой кривой. Она может быть вычислена по формуле*

$$k = \frac{k_n}{\cos\theta}.$$

Доказательство. Пусть кривая C_0 параметризована вектор-функцией $\boldsymbol{\rho}(\tau)$, причем $P = \boldsymbol{\rho}(\tau_0)$, а соответствующие внутренние уравнения имеют вид $u = \rho_1(\tau)$, $v = \rho_2(\tau)$. Так как в точке P у кривых C и C_0 общая касательная, то $\boldsymbol{\varphi}'(t_0) = a\boldsymbol{\rho}'(\tau_0)$ для некоторого $a \neq 0$. Следовательно, $\varphi_1'(t_0) = a\rho_1'(\tau_0)$, $\varphi_2'(t_0) = a\rho_2'(\tau_0)$. Тогда в точке P имеем окончательно

$$k\,\cos\theta = \frac{\mathrm{II}(\varphi_1',\ \varphi_2')}{\mathrm{I}(\varphi_1',\ \varphi_2')} = \frac{\mathrm{II}(a\rho_1',\ a\rho_2')}{\mathrm{I}(a\rho_1',\ a\rho_2')} = \frac{a^2\mathrm{II}(\rho_1',\ \rho_2')}{a^2\mathrm{I}(\rho_1',\ \rho_2')} = \\ = \frac{\mathrm{II}(\rho_1',\ \rho_2')}{\mathrm{I}(\rho_1',\ \rho_2')} = k_0 \cdot \cos\theta_0 = k_n.$$

Что и требовалось доказать. □

Нормальную кривизну k_n поверхности Φ в направлении касательной l будем также называть *нормальной кривизной кривой C в точке P* и обозначать так же: k_n.

Если две кривые на поверхности Φ, проходящие через точку P, имеют в ней общую касательную, то очевидно их нормальные кривизны в этой точке совпадают.

Если $\boldsymbol{k}$ — вектор кривизны кривой C в точке P, то его скалярное произведение на вектор нормали $\boldsymbol{n}$ равно нормальной кривизне кривой C в точке P:

$$\boldsymbol{k} \cdot \boldsymbol{n} = |\boldsymbol{k}| \cdot \cos\theta = k \cdot \cos\theta = k_n.$$

Проиллюстрируем понятие нормальной кривизны.

Теорема 2. *Если две кривые на поверхности* Φ *проходят через точку* P *и имеют в ней общую соприкасающуюся плоскость, не совпадающую с касательной плоскостью* $T_P\Phi$, *то их кривизны в этой точке равны.*

Доказательство. Путь две кривые C_1 и C_2 проходят через точку P. Очевидно, что прямая, по которой пересекается с касательной плоскостью $T_P\Phi$ их общая соприкасающаяся плоскость, является общей касательной этих кривых. Следовательно, их нормальные кривизны в точке P совпадают:

$$k_n(C_1) = k_n(C_2).$$

Если θ — угол между соприкасающейся плоскостью этих кривых и нормалью к поверхности в точке P, то, по теореме Мёнье,

$$k(C_1) = \frac{|k_n(C_1)|}{\cos\theta} = \frac{|k_n(C_2)|}{\cos\theta} = k(C_2),$$

где $k(C_1)$, $k(C_2)$ — кривизны кривых C_1 и C_2 в точке P. Теорема 2 доказана. □

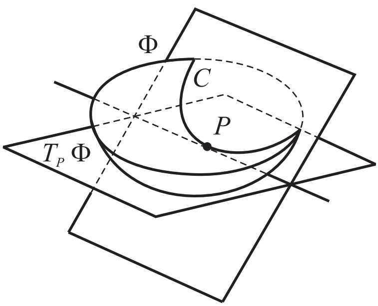


Рис. 60.

Следствие. *Кривизна кривой, лежащей на поверхности, в каждой точке равна кривизне сечения поверхности соприкасающейся плоскостью кривой в этой же точке* (рис. 60). □

§ 7. Соприкасающийся параболоид

Здесь мы займемся локальным описанием поверхности с точностью до бесконечно малых второго порядка.

Пусть P — произвольная точка гладкой поверхности Φ. Для изучения нормальных кривизн в точке P перейдем к явному заданию поверхности Φ. Рассмотрим декартову систему координат xyz, начало

которой совпадает с точкой P, а ось z направлена по нормали к касательной плоскости $T_P\Phi$. Тогда, как показано в § 4, некоторая достаточно малая окрестность точки P на поверхности Φ обладает явным заданием

$$z = f(x,\ y),$$

где функция $f(x,\ y)$ определена и непрерывно дифференцируема в некоторой окрестности точки $(0,\ 0)$ на плоскости xy (рис. 61), причем

$$f(0,\ 0) = f_x(0,\ 0) = f_y(0,\ 0) = 0.$$

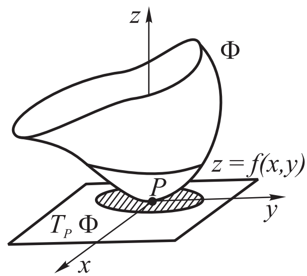


Рис. 61

Оказывается, что за счет поворота осей x и y можно еще более упростить ситуацию и добиться того, чтобы в новых координатах $\widetilde{x}$ и $\widetilde{y}$ в точке $(0,\ 0)$ обращались бы в нуль не только функция f и ее первые частные производные, но и смешанная частная производная:

$$f_{\widetilde{x}\widetilde{y}}(0,\ 0) = 0.$$

Вместо того чтобы доказывать это непосредственно, мы воспользуемся известными результатами о поверхностях второго порядка.

Рассмотрим функцию $\overline{f}(x,\ y)$, определенную по формуле

$$\overline{f}(x,\ y) = \frac{1}{2}(f_{xx}(0,\ 0)x^2 + 2f_{xy}(0,\ 0)xy + f_{yy}(0,\ 0)y^2).$$

Это многочлен второй степени, причем все частные производные первого и второго порядка у $\overline{f}$ и у f в начале координат равны. Нетрудно убедиться, что это выполняется и в любой другой системе координат (с тем же началом).

Функция $\overline{f}$ допускает важное геометрическое истолкование: число $\overline{f}(x_0,\ y_0)$ равняется половине произведения нормальной кривизны поверхности Φ в точке $P(0,\ 0,\ 0)$ в направлении прямой $xy_0 - yx_0 = 0$ на квадрат расстояния точки $(x_0,\ y_0)$ от точки P. Это можно увидеть из формулы для нормальной кривизны и того, что коэффициенты первой и второй квадратичных форм в точке P в параметризации $(u,\ v,\ f(u,\ v))$ вычисляются по очевидным формулам

$$E = 1, \quad F = 0, \quad G = 1;$$

$$L = f_{xx}(0,\ 0), \quad M = f_{xy}(0,\ 0), \quad N = f_{yy}(0,\ 0).$$

Поверхность F, заданная уравнением

$$z = \overline{f}(x,\ y),$$

называется *соприкасающимся параболоидом* поверхности Φ в точке P. Фактически F может представлять собой эллиптический или гиперболический параболоид, параболический цилиндр или плоскость (рис. 62).

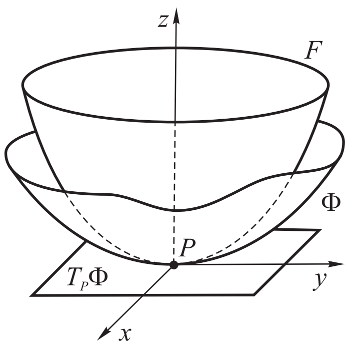


Рис. 62

Соприкасающийся параболоид F обладает рядом замечательных свойств, каждое из которых можно было бы при его определении взять за основу.

Во-первых, как мы уже знаем, F является графиком половины от второй квадратичной формы II в точке P, если II рассматривать как функцию на касательной плоскости $T_P\Phi$.

Во-вторых, в точке P у поверхностей F и Φ нормальные кривизны в любом направлении совпадают.

Наконец, при помощи формулы Тэйлора можно показать, что соприкасающийся параболоид F с точностью до малых второго порядка приближает поверхность Φ в точке P. (Здесь мы условно причисляем плоскость и параболический цилиндр к параболоидам.)

Более точно, пусть Q — точка поверхности Φ, близкая к точке P, а Q' — такая точка параболоида F, что прямая $Q'Q$ перпендикулярна плоскости $T_P\Phi$. Тогда *при стремлении точки Q к точке P отношение квадрата расстояния от Q' до Q к расстоянию от Q до P стремится к нулю* (рис. 63):

$$\lim_{Q \to P} \frac{Q'Q}{PQ^2} = 0.$$

Соприкасающийся параболоид — единственный, обладающий этим свойством.

Типы точек на поверхности. Можно произвести классификацию точек поверхности в соответствии с типом соприкасающегося параболоида. Пусть по-прежнему F — соприкасающийся параболоид поверхности Φ в точке P.

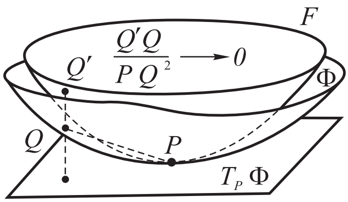


Рис. 63

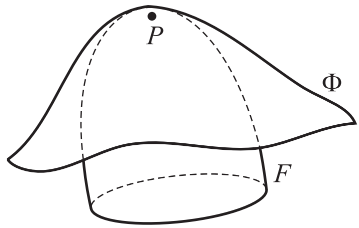


Рис. 64

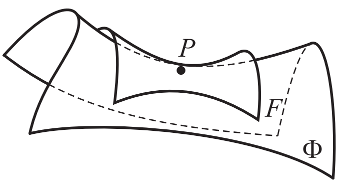


Рис. 65

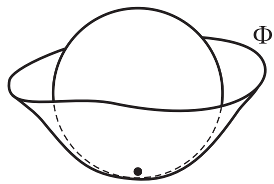


Рис. 66

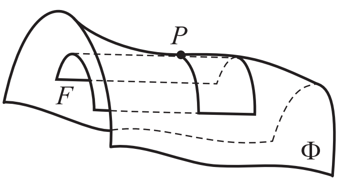


Рис. 67

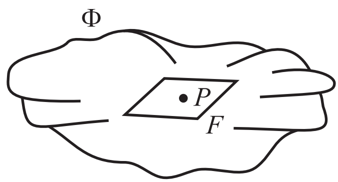
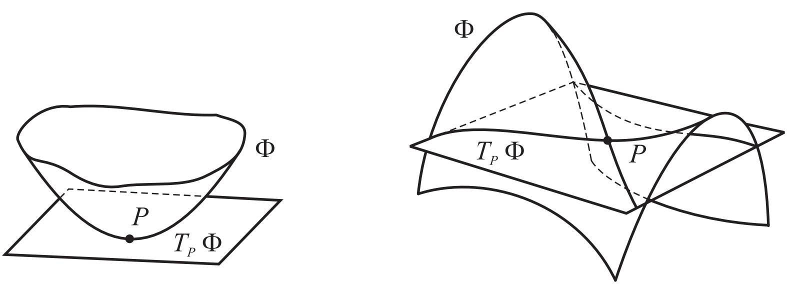


Рис. 68



а *б*

Рис. 69

P называется точкой *эллиптического типа,* если F — эллиптический параболоид (рис. 64). P называется точкой *гиперболического типа,* если F — гиперболический параболоид (рис. 65). P называется точкой *параболического типа*, если F — параболический цилиндр или плоскость (рис. 67, 68). Точка P эллиптического типа называется *точкой округления*, или *омбилической*, если F — параболоид вращения (рис. 66). Наконец, точка P параболического типа называется *точкой уплощения*, если F — плоскость.

Можно показать, что в окрестности точки эллиптического типа поверхность лежит по одну сторону от своей касательной плоскости (рис. 69, a), а в окрестности точки гиперболического типа поверхность пересекается с касательной плоскостью по двум кривым (рис. 69, *б*). В окрестности точки параболического типа ничего определенного про расположение поверхности относительно касательной плоскости заранее сказать нельзя.

§ 8. Главные кривизны и формула Эйлера

Пусть F — соприкасающийся параболоид поверхности Φ в точке P. Из аналитической геометрии известно, что поворотом осей уравнение параболоида F можно привести к виду

$$z = \frac{1}{2}(k_1 x^2 + k_2 y^2).$$

Рассмотрим касательную прямую l, проходящую через точки P и $(x_0,\ y_0)$. Нормальную кривизну поверхности Φ в точке P в направлении прямой l можно вычислить по формуле (см. § 6)

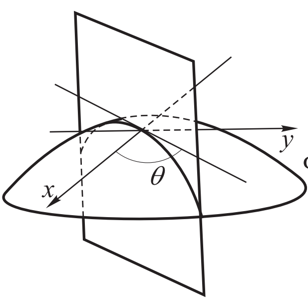


Рис. 70

$$k_n = \frac{\mathrm{II}(x_0,\ y_0)}{\mathrm{I}(x_0,\ y_0)} = \frac{k_1 x_0^2 + k_2 y_0^2}{x_0^2 + y_0^2} = \\ = k_1 \frac{x_0^2}{x_0^2 + y_0^2} + k_2 \frac{y_0^2}{x_0^2 + y_0^2}.$$

Если обозначить через θ угол между осью x и прямой l, то формулу можно записать в виде (рис. 70):

$$k_n(\theta) = k_1 \cos^2 \theta + k_2 \sin^2 \theta.$$

Из этой формулы следует, что если $k_1 < k_2$, то

$$k_1 = k_1 \cos^2 \theta + k_1 \sin^2 \theta \leqslant k_n(\theta) \leqslant k_2 \cos^2 \theta + k_2 \sin^2 \theta = k_2,$$

причем равенство в одном из неравенств может иметь место только в том случае, когда $\cos\theta = 0$ или $\sin\theta = 0$, т. е. когда прямая l совпадает с одной из координатных осей. Если же $k_1 = k_2$, то $k_n(\theta) \equiv k_1$.

Таким образом, мы доказали следующую теорему Эйлера.

Теорема 1 (Эйлер). *В каждой точке гладкой поверхности существуют две перпендикулярные касательные прямые l_1 и l_2, в направлении которых нормальная кривизна поверхности принимает наибольшее и наименьшее значения k_1 и k_2. Если l — произвольная касательная прямая, образующая угол θ с прямой l_1, то нормальная кривизна в направлении l вычисляется по формуле Эйлера*

$$k_n = k_1 \cos^2 \theta + k_2 \sin^2 \theta. \quad \square$$

Наибольшую и наименьшую нормальные кривизны поверхности в точке P называют ее *главными кривизнами* в этой точке, а направления, в которых достигаются главные кривизны, называются *главными направлениями.* Как мы видели, если главные кривизны в точке P различны: $k_1 \neq k_2$, то главные направления определены однозначно и они перпендикулярны друг другу. Если же главные кривизны равны: $k_1 = k_2$, то любое направление будет главным.

Гауссова кривизна и средняя кривизна. Произведение главных кривизн k_1, k_2 обозначается через K и называется *гауссовой кривизной* поверхности в точке P:

$$K = k_1 \cdot k_2.$$

В то время как сам соприкасающийся параболоид F определяется главными кривизнами k_1, k_2 и главными направлениями l_1, l_2, его *тип* полностью определяется *произведением* главных кривизн $k_1 k_2 = K$:

если $k_1 k_2 > 0$, то F — эллиптический параболоид;

если $k_1 k_2 < 0$, то F — гиперболический параболоид;

если $k_1 k_2 = 0$, то F — параболический цилиндр или плоскость.

Таким образом, тип точки P на поверхности полностью определяется гауссовой кривизной K поверхности в этой точке:

если $K > 0$, то P — точка эллиптического типа (рис. 71, 72);

если $K < 0$, то P — точка гиперболического типа (рис. 73);

если $K = 0$, то P — точка параболического типа (рис. 74).

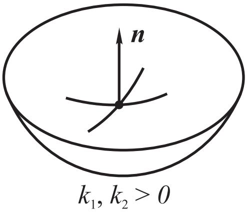


Рис. 71

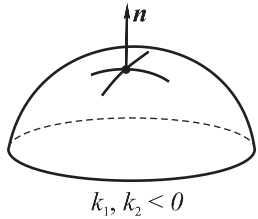


Рис. 72

Полусумма главных кривизн k_1, k_2 обозначается через H и называется *средней кривизной* поверхности в точке P:

$$H = \frac{k_1 + k_2}{2}.$$

(Из формулы Эйлера следует, что число H равняется нормальной кривизне в направлении биссектрисы угла, образованного главными направлениями.)

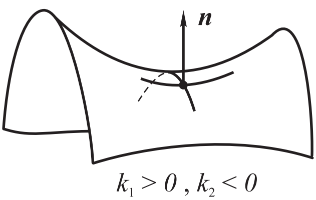


Рис. 73

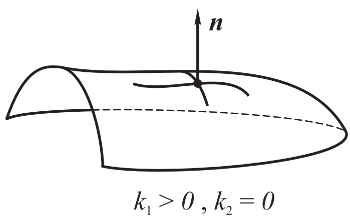


Рис. 74

Главные кривизны k_1, k_2 в точке P полностью определяются гауссовой кривизной K и средней кривизной H: по теореме Виета k_1 и k_2 являются корнями уравнения

$$x^2 - 2Hx + K = 0.$$

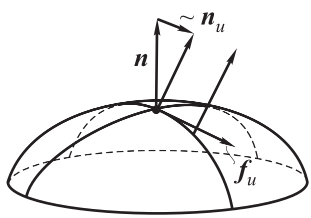


Рис. 75

Заметим еще, что P является точкой округления тогда и только тогда, когда $H = K$, а точкой уплощения — тогда и только тогда, когда $H = K = 0$.

Наконец, отметим одно важное свойство главных направлений. Мы

сформулируем его в том виде, в котором оно будет нами использоваться в § 11.

Теорема 2 (Родриг). *Пусть $\boldsymbol{f}(u,\ v)$ — параметризация гладкой поверхности Φ, и $P = \boldsymbol{f}(u_0,\ v_0)$ — точка на поверхности. Если направления координатных линий в точке P являются главными, то выполняются равенства*

$$\boldsymbol{n}_u(u_0,\ v_0) = -k_1\boldsymbol{f}_u(u_0,\ v_0),$$
$$\boldsymbol{n}_v(u_0,\ v_0) = -k_2\boldsymbol{f}_v(u_0,\ v_0),$$

где k_1 и k_2 — главные кривизны в точке P (рис. 75).

Доказательство. Нетрудно видеть, что справедливость доказываемых равенств зависит не от конкретного выбора системы внутренних координат u, v, а только от направлений координатных линий в точке P. Поэтому достаточно проверить доказываемые равенства непосредственно, воспользовавшись для этого, например, явным заданием поверхности

$$z = \varphi(x,\ y),$$

при котором

$$\varphi(0,\ 0) = \varphi_x(0,\ 0) = \varphi_y(0,\ 0) = \varphi_{xy}(0,\ 0) = 0,$$

а точка P имеет координаты 0, 0, 0. Соответствующая параметризация имеет вид $\boldsymbol{f}\,(u,\ v) = (u,\ v,\ \varphi(u,\ v))$. Рассмотрим вектор-функцию $\boldsymbol{N}(u,\ v) = \boldsymbol{f}_u(u,\ v) \times \boldsymbol{f}_v(u,\ v)$. Тогда

$$\boldsymbol{n}(u,\ v) = \frac{\boldsymbol{N}(u,\ v)}{|\boldsymbol{N}(u,\ v)|},$$
$$\boldsymbol{n}_u(u,\ v) = \frac{\boldsymbol{N}_u(u,\ v)\cdot|\boldsymbol{N}(u,\ v)| - \boldsymbol{N}(u,\ v)\cdot\partial_u|\boldsymbol{N}(u,\ v)|}{|\boldsymbol{N}(u,\ v)|^2}$$

Поскольку $\boldsymbol{f}_u(u,\ v) = (1,\ 0,\ \varphi_x(u,\ v))$, $\boldsymbol{f}_v(u,\ v) = (0,\ 1,\ \varphi_y(u,\ v))$, то нетрудно видеть, что

$$\boldsymbol{N}(u,\ v) = (-\varphi_x,\ -\varphi_y,\ 1),$$
$$\boldsymbol{N}_u(u,\ v) = (-\varphi_{xx},\ 0,\ 0),$$
$$|\boldsymbol{N}(u,\ v)| = \sqrt{\varphi_x^2 + \varphi_y^2 + 1},$$
$$\partial_u|\boldsymbol{N}(u,\ v)| = \frac{\varphi_x\cdot\varphi_{xx} + \varphi_y\cdot\varphi_{yy}}{\sqrt{\varphi_x^2 + \varphi_y^2 + 1}},$$

где для краткости опущены аргументы u, v у функций φ_x, φ_y, φ_{xx}, φ_{yy}. Далее, ясно, что

$$\boldsymbol{N}(0,\ 0) = (0,\ 0,\ 1),$$
$$\boldsymbol{N}_n(0,\ 0) = (-k_1,\ 0,\ 0), \quad |\boldsymbol{N}(0,\ 0)| = 1,$$

и что функция $\partial_u(\boldsymbol{N}(u,\ v))$ в точке $(0, 0)$ обращается в нуль. Подставляя, получаем требуемое равенство:

$$\boldsymbol{n}_u(0,\ 0) = \frac{(-k_1,\ 0,\ 0)\cdot 1 - (0,\ 0,\ 1)\cdot 0}{1^2} = (-k_1,\ 0,\ 0) = \\ = -k_1 \cdot \boldsymbol{f}_u(0,\ 0).$$

Равенство $\boldsymbol{n}_v(0,\ 0) = -k_2\boldsymbol{f}_v(0,\ 0)$ проверяется аналогично. □

§ 9. Нахождение главных направлений и главных кривизн

Пусть Φ — гладкая поверхность, заданная векторным уравнением $\boldsymbol{r} = \boldsymbol{f}\,(u,\ v)$, а $P = \boldsymbol{f}\,(u_0,\ v_0)$ — точка на ней. Рассмотрим в касательной плоскости $T_P\Phi$ аффинную систему координат ξ, η, в которой векторы $\boldsymbol{f}_u(0,\ 0)$ и $\boldsymbol{f}_v(0,\ 0)$ будут направляющими векторами осей ξ и η. В такой системе точка с радиус-вектором $\boldsymbol{f}\,(u_0,\ v_0) + \xi_0\boldsymbol{f}_u(0,\ 0) + \eta_0\boldsymbol{f}_v(0,\ 0)$ будет иметь координаты ξ_0, η_0. Будем рассматривать первую и вторую квадратичную формы как функции на плоскости $T_P\Phi$:

$$\mathrm{I}(\xi,\ \eta) = E\xi^2 + 2F\xi\eta + G\eta^2,$$
$$\mathrm{II}(\xi,\ \eta) = L\xi^2 + 2M\xi\eta + N\eta^2.$$

Тогда нормальная кривизна поверхности Φ в точке P в направлении прямой, проходящей через точку P и точку с координатами $(\xi_0,\ \eta_0)$, вычисляется по формуле

$$k(\xi_0,\ \eta_0) = \frac{\mathrm{II}(\xi_0,\ \eta_0)}{\mathrm{I}(\xi_0,\ \eta_0)}.$$

Главные направления. Следовательно, весь вопрос о нахождении главных направлений сводится к нахождению точек максимума и минимума функции $k(\xi,\ \eta)$.

Пусть $k(\xi_0,\ \eta_0)$ — главная кривизна, причем $\eta_0 \neq 0$. Положим $t_0 = \xi_0/\eta_0$. Тогда

$$k(\xi_0,\ \eta_0) = \frac{\mathrm{II}(\xi_0,\ \eta_0)}{\mathrm{I}(\xi_0,\ \eta_0)} = \frac{Et_0^2 + 2Ft_0 + G}{Lt_0^2 + 2Mt_0 + N}.$$

Введем обозначения:

$$\varphi_1(t) = Et^2 + 2Ft + G, \quad \varphi_2(t) = Lt^2 + 2Mt + N.$$

Тогда

$$k(\xi_0,\ \eta_0) = \frac{\varphi_2(t_0)}{\varphi_1(t_0)}.$$

Поскольку, по предположению, функция $\varphi_2(t)/\varphi_1(t)$ достигает в точке t_0 максимума или минимума, то ее производная должна обращаться в этой точке в нуль. Так как

$$\left(\frac{\varphi_2(t)}{\varphi_1(t)}\right)' = \frac{\varphi_2'(t)\varphi_1(t) - \varphi_1'(t)\varphi_2(t)}{\varphi_1^2(t)},$$

то, приравнивая числитель нулю, получаем

$$(2Lt + 2M)(Et^2 + 2Ft + G) - (2Et + 2F)(Lt^2 + 2Mt + N) = 0.$$

Сокращая на 2, раскрывая скобки и приводя подобные члены, получим окончательно

$$t_0^2(LF - ME) + t_0(LG - NE) + (MG - NF) = 0.$$

Если $\xi_0 \neq 0$, то мы можем положить $s_0 = \eta_0/\xi_0$ и рассмотреть функции

$$\psi_1(s) = E + 2Fs + Gs^2, \quad \psi_2(s) = L + 2Ms + Ns^2.$$

Тогда $k(\xi_0,\ \eta_0) = \psi_2(s_0)/\psi_1(s_0)$.

Так же как в случае функций φ_1 и φ_2, получаем, что s_0 является корнем уравнения $\psi_2'(s)\psi_1(s) - \psi_1'(s) \times \psi_2(s) = 0$. Подставляя вместо $\psi_1,\ \psi_2,\ \psi_1',\ \psi_2'$ их выражения через s, деля на 2, раскрывая скобки и приводя подобные члены, получим

$$(ME - LF) + s_0(NE - LG) + (NF - MG)s_0^2 = 0.$$

Если первое уравнение домножить на η_0^2 или второе домножить на $-\xi_0^2$, то получится однородное уравнение относительно ξ_0 и η_0:

$$\xi_0^2(LE - ME) + \xi_0\eta_0(LG - NE) + \eta_0^2(MG - NF) = 0.$$

Его удобно записывать при помощи определителя:

$$\begin{vmatrix} -\eta_0^2 & \xi_0\eta_0 & -\xi_0^2 \\ E & F & G \\ L & M & N \end{vmatrix} = 0.$$

Таким образом, *решена задача отыскания главных направлений.*

Главные кривизны. Чтобы найти главные кривизны, можно было бы теперь решить одно из полученных уравнений, а затем найти кривизну по общей формуле. Мы поступим иначе.

Предположим, что мы хотим найти направление, в котором нормальная кривизна принимает данное значение k. Для этого нам надо было бы решить уравнение

$$\mathrm{II}(\xi,\ \eta) = \mathrm{I}(\xi,\ \eta)k,$$

или, что равносильно,

$$(L - kE)\xi^2 + 2(M - kF)\xi\eta + (N - kG)\eta^2 = 0.$$

Это уравнение можно рассматривать как квадратное, например, относительно ξ/η. Если k — главная кривизна, то у него есть единственное решение, и, следовательно, дискриминант равен нулю:

$$(L - kE)(N - kG) - (M - kF)^2 = 0.$$

Отсюда получаем квадратное уравнение для нахождения главных кривизн k_1, k_2 в точке P:

$$(EG - F^2)k^2 + (2MF - EN - LG)k + (LN - M^2) = 0.$$

Кроме того, из этого уравнения, в силу формул Виета, получаем важное следствие — выражение гауссовой и средней кривизн через коэффициенты первой и второй квадратичных форм поверхности в данной точке:

$$K = k_1k_2 = \frac{LN - M^2}{EG - F^2},$$
$$2H = k_1 + k_2 = \frac{EN - 2MF + LG}{EG - F^2}.$$

§ 10. Площадь поверхности

Дадим определение площади гладкой поверхности, исходя из того, что определение площади нам известно для замкнутых плоских областей, ограниченных конечным числом гладких кривых.

Пусть Φ — гладкая поверхность, заданная уравнением $\boldsymbol{r} = \boldsymbol{f}\ (u,\ v)$, где $(u,\ v) \in V \subset \mathbf{R}^2$. Возьмем в области V некоторую замкнутую

область $\widetilde{D}$, ограниченную конечным числом гладких кривых. Чтобы определить площадь ее образа — области $D = \boldsymbol{f}\,(\widetilde{D})$ — поступим так. Разделим область $\widetilde{D}$ на достаточно маленькие (по диаметру) замкнутые области $\widetilde{D}_i$ при помощи кусочно гладких кривых. Соответственно и область D на поверхности разобьется на области D_i, ограниченные кусочно гладкими кривыми. В каждой области D_i выберем произвольную точку P_i. В ней проведем касательную плоскость $T_{P_i}\Phi$ и область D_i спроектируем на эту плоскость. Если размеры областей $\widetilde{D}_i$ достаточно малы, то при этом область D_i взаимно однозначно отобразится на некоторую *плоскую* область G_i, площадь которой нам известна. Кажется правдоподобным, что площадь $S(D_i)$ должна «мало отличаться» от площади $S(G_i)$, а площадь всей поверхности должна «мало отличаться» от суммы площадей всех областей G_i (рис. 76).

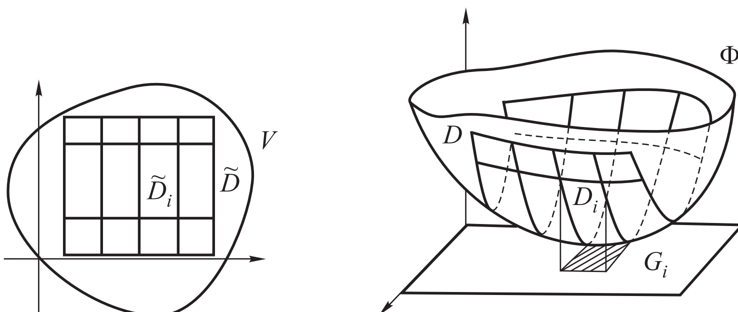


Рис. 76

Если при неограниченном уменьшении (по диаметру) областей D_i сумма площадей областей G_i будет стремиться к некоторому пределу, то этот предел называется *площадью* области D:

$$S(D) = \lim_{\max d(D_i)\to 0} \left(\sum_i S(G_i) \right).$$

Теорема. *Всякая замкнутая и ограниченная область D с кусочно гладкой границей на поверхности Φ имеет определенную площадь.*

Эту площадь можно вычислить по формуле

$$S(D) = \iint\limits_{\widetilde{D}} \sqrt{EG - F^2}\, du\, dv.$$

Доказательство. В силу аддитивности левой и правой частей доказываемого равенства достаточно ограничиться случаем, когда область D настолько мала, что, например, допускает явное задание

$$z = \varphi(x,\ y),$$

где φ — гладкая функция, заданная в области G. Тогда, как известно из анализа, площадь области D можно найти по формуле

$$S(D) = \iint\limits_{G} \sqrt{1 + \varphi_x^2 + \varphi_y^2}\, dx\, dy.$$

Пусть переход от явного задания $z = \varphi(x,\ y)$ к параметрическому $\boldsymbol{r} = \boldsymbol{f}\,(u,\ v)$ осуществляется при помощи замены переменных

$$x = \xi(u,\ v), \quad y = \eta(u,\ v).$$

Тогда, воспользовавшись формулой замены переменных под знаком двойного интеграла, мы получаем:

$$S(D) = \iint\limits_{G} \sqrt{1 + \varphi_x^2 + \varphi_y^2}\, dx\, dy = \\ = \iint\limits_{\widetilde{D}} \sqrt{1 + \varphi_x^2(\xi,\ \eta) + \varphi_y^2(\xi,\ \eta)}\, J\, du\, dv,$$

где $J = J(u,\ v) = \begin{vmatrix} \xi_u(u,\ v) & \eta_u(u,\ v) \\ \xi_v(u,\ v) & \eta_v(u,\ v) \end{vmatrix}$ — якобиан замены. Поскольку, очевидно,

$$\boldsymbol{f}\,(u,\ v) = (\xi(u,\ v),\ \eta(u,\ v),\ \varphi(\xi(u,\ v),\ \eta(u,\ v))),$$

то

$$\boldsymbol{f}_u(u,\ v) = (\xi_u,\ \eta_u,\ \varphi_x\xi_u + \varphi_y\eta_u),$$
$$\boldsymbol{f}_v(u,\ v) = (\xi_v,\ \eta_v,\ \varphi_x\xi_v + \varphi_y\eta_v),$$

и, как нетрудно проверить,

$$\boldsymbol{f}_u \times \boldsymbol{f}_v = J \cdot (-\varphi_x,\ -\varphi_y,\ 1).$$

Отсюда $\sqrt{EG - F^2} = |\boldsymbol{f}_u \times \boldsymbol{f}_v| = \sqrt{1 + \varphi_x^2 + \varphi_y^2}\,|J|$, т. е. наш интеграл совпадает с указанным в формулировке теоремы. □

§ 11. Сферическое отображение поверхности

Для изучения искривленности поверхностей очень полезным оказывается некоторое их отображение в единичную сферу, которое мы сейчас опишем.

Пусть Φ — гладкая поверхность, P — произвольная ее точка. Пусть $\boldsymbol{n}$ — единичный вектор нормали к поверхности в точке P. Отложим вектор $\boldsymbol{n}$ из начала координат. Тогда его конец определит некоторую точку $\Gamma(P)$ единичной сферы $S^2 \subset \mathbf{R}^3$. Построенное отображение Γ

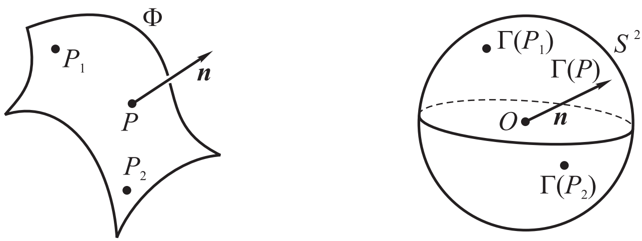


Рис. 77

поверхности Φ в единичную сферу S^2 называется *сферическим* или *гауссовым* отображением (рис. 77):

$$\Gamma\colon \quad \Phi \to S^2.$$

(Стоит отметить, что при сферическом отображении касательная плоскость к поверхности в точке P параллельна касательной плоскости к сфере в точке $\Gamma(P)$.) Образы точек и множеств при сферическом отображении называются их *сферическими изображениями* (рис. 78).

Примеры. 1. Если Φ — область на сфере S^2, то ее сферическое отображение есть тождественное, а сферическое изображение совпадает с Φ.

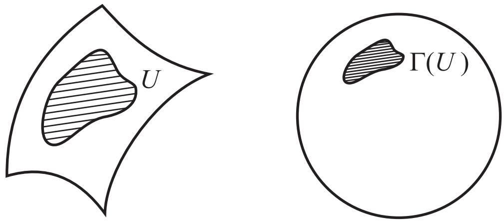


Рис. 78

2. Сферическое изображение любой образующей на цилиндрической поверхности есть точка, а сферическое изображение всей поверхности есть некоторая дуга большого круга на сфере.

3. Сферическое отображение плоской области есть постоянное отображение, а ее сферическое изображение есть точка (рис. 79).

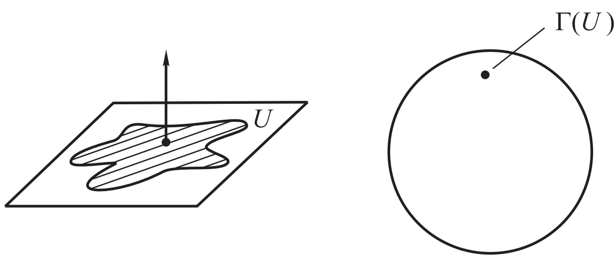


Рис. 79

4. Сферическое изображение эллиптического или гиперболического параболоида есть открытая полусфера (рис. 80).

Из теоремы об обратной функции и теоремы Родрига нетрудно вывести, что если P — точка эллиптического или гиперболического типа, то в достаточно малой окрестности точки P сферическое отображение взаимно однозначно. Это означает, что в такой окрестности не найдется двух точек, нормали к поверхности в которых были бы друг другу параллельны.

При помощи сферического отображения можно дать геометрическую интерпретацию гауссовой кривизны.

Теорема. *Пусть U — малая окрестность точки P на поверхности Φ. Тогда при стремлении диаметра области U к нулю (другими словами, при стягивании ее к точке P) отношение площади ее сфе-*

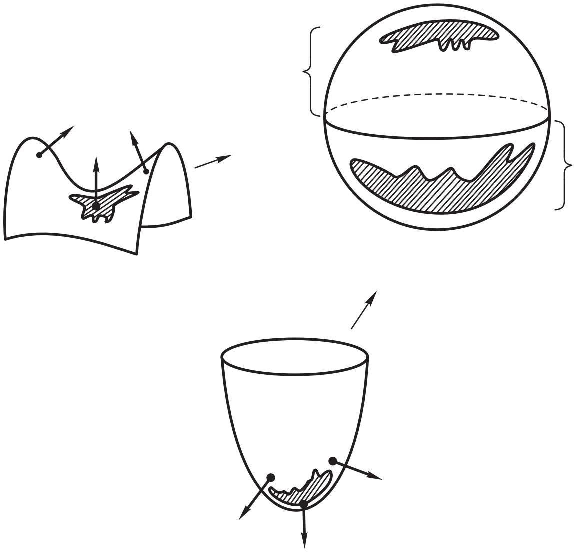


Рис. 80

рического изображения $\Gamma(U)$ *к площади самой окрестности* U *стремится к абсолютной величине гауссовой кривизны поверхности* Φ *в точке* P*:*

$$\lim_{d(U)\to 0} \frac{S(\Gamma(U))}{S(U)} = |K(P)|.$$

Доказательство. Мы ограничимся случаем, когда точка P принадлежит к эллиптическому или гиперболическому типу. В этом случае отображение Γ взаимно однозначно на U и имеет место равенство

$$S(\Gamma(U)) = \iint\limits_W |\boldsymbol{n}_u(u,\ v) \times \boldsymbol{n}_v(u,\ v)|\, du\, dv,$$

где W — координатная окрестность, соответствующая окрестности U: $W = \boldsymbol{f}^{-1}(U)$ (где $\boldsymbol{f}$ параметризует Φ). С другой стороны, всегда

$$S(U) = \iint\limits_W |\boldsymbol{f}_u(u,\ v) \times \boldsymbol{f}_v(u,\ v)|\, du\, dv.$$

По теореме о среднем значении получаем

$$S(\Gamma(U)) = \iint\limits_W |\boldsymbol{n}_u \times \boldsymbol{n}_v|\, du\, dv = |\boldsymbol{n}_u(u',\ v') \times \boldsymbol{n}_v(u',\ v')| \cdot S(W),$$

$$S(U) = \iint\limits_W |\boldsymbol{f}_u \times \boldsymbol{f}_v|\, du\, dv = |\boldsymbol{f}_u(u'',\ v'') \times \boldsymbol{f}_v(u'',\ v'')| \cdot S(W),$$

где $(u',\ v') \in W$, $(u'',\ v'') \in W$ — некоторые точки. При стягивании области U к точке P эти точки стремятся к $(u_0,\ v_0)$. Поэтому

$$\lim_{d(U)\to 0} \frac{S(\Gamma(U))}{S(U)} = \lim_{\substack{(u',v')\to(u_0,v_0)\\(u'',v'')\to(u_0,v_0)}} \frac{|\boldsymbol{n}_u(u',v') \times \boldsymbol{n}_v(u',v')|}{|\boldsymbol{f}_u(u'',v'') \times \boldsymbol{f}_v(u'',v'')|} =$$

$$= \frac{|\boldsymbol{n}_u(u_0,v_0) \times \boldsymbol{n}_v(u_0,v_0)|}{|\boldsymbol{f}_u(u_0,\ v_0) \times \boldsymbol{f}_v(u_0,\ v_0)|}.$$

Значение полученного выражения проще всего подсчитать, воспользовавшись теоремой Родрига. Если внутренние координаты выбраны так, что координатные линии проходят через точку P в главных направлениях, то

$$\boldsymbol{n}_u(u_0,\ v_0) = -k_1\boldsymbol{f}_u(u_0,\ v_0), \quad \boldsymbol{n}_v(u_0,\ v_0) = -k_2\boldsymbol{f}_u(u_0,\ v_0),$$

где k_1 и k_2 — главные кривизны в точке P. Поэтому

$$\frac{|\boldsymbol{n}_u(u_0,\ v_0) \times \boldsymbol{n}_v(u_0,\ \ v_0)|}{|\boldsymbol{f}_u(u_0,\ v_0) \times \boldsymbol{f}_v(u_0,\ v_0)|} = |k_1 \cdot k_2| = |K(P)|,$$

что и требовалось доказать. □

§ 12. Внутренняя геометрия поверхности

Определение. Пусть даны две поверхности — Φ и $\widetilde{\Phi}$. Предположим, что задано непрерывное биективное отображение i: $\Phi \to \widetilde{\Phi}$ одной из них в другую. Такое отображение устанавливает взаимно однозначное соответствие между точками обеих поверхностей. При этом соответствии каждой кривой C на поверхности Φ отвечает некоторая кривая $\widetilde{C}$ на поверхности $\widetilde{\Phi}$, и наоборот:

$$\widetilde{C} = i(C), \quad C = i^{-1}(\widetilde{C}).$$

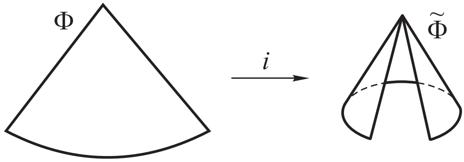


Рис. 81

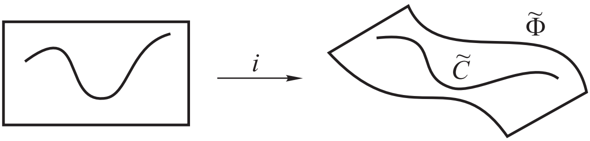


Рис. 82

Если при этом длина каждой кривой C равна длине соответствующей кривой $\widetilde{C}$, то говорят, что $\widetilde{\Phi}$ получается из Φ при помощи изгибания, а само отображение i называется *изгибанием*, или *изометрией* (рис. 81, 82).

Внутренняя геометрия поверхности изучает те свойства поверхностей и фигур на них, которые не меняются при изгибаниях. С этой точки зрения вся планиметрия представляет собой внутреннюю геометрию плоскости. Это вполне согласуется с тем обычным представлением, что в планиметрии изучаются свойства фигур, не меняющиеся при изометриях плоскости, т. е. наложениях.

Главным для нас будет следующий пример изгибания. Пусть поверхности Φ и $\widetilde{\Phi}$ параметризуются вектор-функциями $\boldsymbol{f}\,(u,\,v)$ и $\widetilde{\boldsymbol{f}}(u,\,v)$, заданными в одной и той же области V, и пусть отображение i ставит в соответствие точке P на поверхности Φ точку $\widetilde{P}$ на поверхности $\widetilde{\Phi}$, имеющую такие же внутренние координаты (рис. 83):

$$i(\boldsymbol{f}\,(u,\,v)) = \widetilde{\boldsymbol{f}}(u,\,v).$$

Если при этом в соответствующих точках будут совпадать коэффициенты первой квадратичной формы поверхностей Φ и $\widetilde{\Phi}$:

$$E(u,\,v) \equiv \widetilde{E}(u,\,v), \quad F(u,\,v) \equiv \widetilde{F}(u,\,v),$$
$$G(u,\,v) \equiv \widetilde{G}(u,\,v),$$

то, как это следует из формулы длины кривой на поверхности, отображение i будет изгибанием. Обратное утверждение тоже вер-

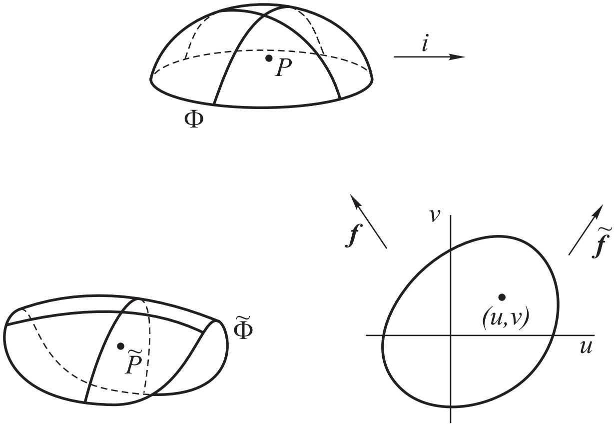


Рис. 83

но. Его доказательство мы оставляем читателю в виде упражнения.

Таким образом, можно сказать, что к внутренней геометрии относятся те свойства и величины, которые могут быть охарактеризованы или вычислены в терминах первой квадратичной формы поверхности. Помимо длин кривых, это углы между кривыми и площади фигур на поверхности.

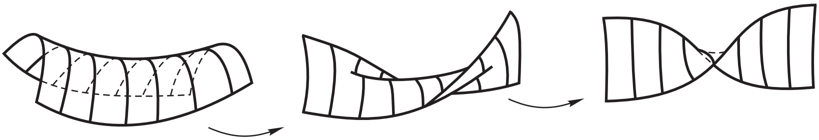


Рис. 84

В следующем параграфе мы докажем, что к внутренней геометрии относится и такая важнейшая характеристика поверхности, как ее гауссова кривизна. Затем мы изучим еще некоторые объекты внутренней геометрии.

Замечание. Точно так же, как наложение (первого рода) плоскости является результатом *перемещения*, изометрию поверхности часто можно представить в виде результата некоторого процесса. Для более точной формулировки введем понятие *непрерывного изгибания*. Пусть поверхности Φ и $\widetilde{\Phi}$ обладают параметризациями $\boldsymbol{f}\,(u,\,v)$ и $\widetilde{\boldsymbol{f}}(u,\,v)$, заданными в одной и той же области V. Пусть задана гладкая вектор-функция трех переменных $\boldsymbol{F}(t,\,u,\,v)$, определенная при $t \in [0,\,1]$ и

$(u, v) \in V$. Пусть для каждого $t_0 \in [0, 1]$ отображение $\boldsymbol{F}(t_0, u, v)$: $V \to \mathbf{R}^3$ является регулярной параметризацией некоторой гладкой поверхности Φ_{t_0}, причем $\boldsymbol{F}(0, u, v) \equiv \boldsymbol{f}(u, v)$ и $\boldsymbol{F}(1, u, v) \equiv \widetilde{\boldsymbol{f}}(u, v)$. (В частности, $\Phi_0 = \Phi$, $\Phi_1 = \widetilde{\Phi}$.) Если при этом для любого $t_0 \in [0, 1]$ отображение поверхности Φ в поверхность Φ_{t_0}, заданное формулой

$$\boldsymbol{f}(u, v) \longmapsto \boldsymbol{F}(t_0, u, v),$$

является изгибанием, то говорят, что $\boldsymbol{F}$ задает *непрерывное изгибание* поверхности Φ в поверхность $\widetilde{\Phi}$. Из определения видно, что в этом случае отображение i поверхности Φ в $\widetilde{\Phi}$, заданное формулой

$$i(\boldsymbol{f}(u, v)) = \boldsymbol{F}(1, u, v) = \widetilde{\boldsymbol{f}}(u, v),$$

является изгибанием (рис. 84).

§ 13. Формула для гауссовой кривизны и следствия из нее. Основные уравнения теории поверхностей

В этом параграфе мы выведем замечательную формулу Гаусса, которая выражает гауссову кривизну поверхности только через коэффициенты первой квадратичной формы этой поверхности и их частные производные первого и второго порядка.

Уже из самого факта существования такой формулы можно сделать несколько очень важных выводов. Первый вывод — инвариантность гауссовой кривизны относительно изгибаний.

Теорема Гаусса (theorema egregium[14]). *Если одна гладкая поверхность получается из другой при помощи изгибания, то гауссовы кривизны этих поверхностей в соответственных точках совпадают. Другими словами, гауссова кривизна поверхности не меняется при изгибании.* □

Таким образом, гауссова кривизна относится к внутренней геометрии поверхностей. Как мы знаем, тип точки на поверхности определяется гауссовой кривизной поверхности в этой точке. Поэтому точку, например, эллиптического типа нельзя превратить изгибанием в точку гиперболического или параболического типа (рис. 85—87).

Другой вывод касается сферического отображения поверхности. Ясно, что при изгибании сферические изображения всей поверхности

[14] egregium (лат.) — славная, отличная, превосходная.

и различных фигур на ней могут меняться. Однако *площадь* сферического изображения фигуры при изгибании не меняется. Это следует из того обстоятельства, что площадь сферического изображения области Ω на поверхности совпадает с абсолютной величиной ее полной кривизны, которая выражается интегралом $\iint\limits_{\Omega} K\,dS$, ср. § 11. А указанный интеграл, в силу теоремы Гаусса, при изгибании не изменяется (рис. 88). (Чтобы эти рассуждения имели силу, необходимо предположить, что сферическое отображение является взаимно однозначным в области Ω и той области $\widetilde{\Omega}$, которая соответствует ей при изгибании.)

Формула Гаусса. Перейдем к выводу формулы Гаусса. Пусть гладкая поверхность Φ задана параметрическим уравнением $\boldsymbol{r} = \boldsymbol{f}\,(u,\,v)$. Тогда ее гауссова кривизна в точке $P = \boldsymbol{f}\,(u,\,v)$, как было показано в § 9, вычисляется по формуле

$$K(u,\,v) = \frac{LN - M^2}{EG - F^2}.$$

(Напомним, что коэффициенты E, F, G, L, M, N первой и второй квадратичной форм являются *функциями* от переменных u, v.) Как мы знаем, для коэффициентов L, M, N имеются формулы

$$L = \frac{(\boldsymbol{f}_{uu},\boldsymbol{f}_{u},\boldsymbol{f}_{v})}{\sqrt{EG-F^2}}, \quad M = \frac{(\boldsymbol{f}_{uv},\boldsymbol{f}_{u},\boldsymbol{f}_{v})}{\sqrt{EG-F^2}}, \quad N = \frac{(\boldsymbol{f}_{vv},\boldsymbol{f}_{u},\boldsymbol{f}_{v})}{\sqrt{EG-F^2}}.$$

Отсюда получаем формулу для гауссовой кривизны в точке P:

$$K(u,\,v) = \frac{(\boldsymbol{f}_{uu},\boldsymbol{f}_{u},\boldsymbol{f}_{v})(\boldsymbol{f}_{vv},\boldsymbol{f}_{u},\boldsymbol{f}_{v}) - (\boldsymbol{f}_{uv},\boldsymbol{f}_{u},\boldsymbol{f}_{v})^2}{(EG-F^2)^2}.$$

Чтобы преобразовать числитель к удобному для нас виду, воспользуемся следующей простой леммой из векторной алгебры.

Лемма. *Пусть даны векторы* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{a}'$, $\boldsymbol{b}'$, $\boldsymbol{c}' \in \mathbf{R}^3$. *Тогда имеет место равенство, связывающее их скалярные и смешанные произведения:*

$$(\boldsymbol{a},\,\boldsymbol{b},\,\boldsymbol{c})(\boldsymbol{a}',\,\boldsymbol{b}',\,\boldsymbol{c}') = \begin{vmatrix} (\mathbf{a}\cdot\mathbf{a}') & (\mathbf{a}\cdot\mathbf{b}') & (\mathbf{a}\cdot\mathbf{c}') \\ (\mathbf{b}\cdot\mathbf{a}') & (\mathbf{b}\cdot\mathbf{b}') & (\mathbf{b}\cdot\mathbf{c}') \\ (\mathbf{c}\cdot\mathbf{a}') & (\mathbf{c}\cdot\mathbf{b}') & (\mathbf{c}\cdot\mathbf{c}') \end{vmatrix}.$$

Доказательство. Проверим эту формулу в координатах. Пусть вектор $\boldsymbol{a}$ имеет координаты a_1, a_2, a_3, вектор $\boldsymbol{a}'$ имеет координаты a_1', a_2', a_3' и т. д. Тогда требуемое равенство получается в результате несложных преобразований

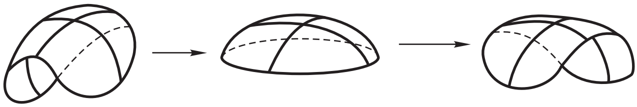


Рис. 85

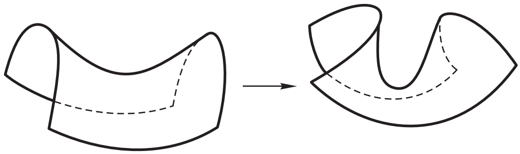


Рис. 86

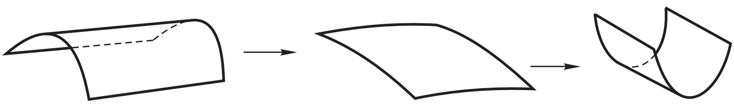


Рис. 87

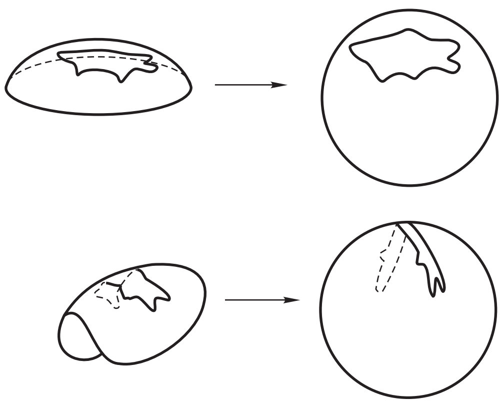


Рис. 88

$$(\boldsymbol{a},\,\boldsymbol{b},\,\boldsymbol{c})(\boldsymbol{a}',\,\boldsymbol{b},\,\boldsymbol{c}') = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \cdot \begin{vmatrix} a_1' & a_2' & a_3' \\ b_1' & b_2' & b_3' \\ c_1' & c_2' & c_3' \end{vmatrix} =$$

$$= \left| \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \cdot \begin{pmatrix} a_1' & b_1' & c_1' \\ a_2' & b_2' & c_2' \\ a_3' & b_3' & c_3' \end{pmatrix} \right| = \begin{vmatrix} (\boldsymbol{a}\cdot\boldsymbol{a}') & \boldsymbol{a}\cdot\boldsymbol{b}') & (\boldsymbol{a}\cdot\boldsymbol{c}') \\ (\boldsymbol{b}\cdot\boldsymbol{a}') & (\boldsymbol{b}\cdot\boldsymbol{b}') & (\boldsymbol{b}\cdot\boldsymbol{c}') \\ (\boldsymbol{c}\cdot\boldsymbol{a}') & (\boldsymbol{c}\cdot\boldsymbol{b}') & (\boldsymbol{c}\cdot\boldsymbol{c}') \end{vmatrix}.$$

Лемма доказана. □

Применяя лемму к выражениям $(\boldsymbol{f}_{uu}, \boldsymbol{f}_u, \boldsymbol{f}_v)(\boldsymbol{f}_{vv}, \boldsymbol{f}_u, \boldsymbol{f}_v)$ и $(\boldsymbol{f}_{uv}, \boldsymbol{f}_u, \boldsymbol{f}_v)^2$, получаем такую довольно громоздкую формулу для гауссовой кривизны:

$$K(u,\,v) = \frac{1}{(EG-F^2)^2}\left\{\begin{vmatrix} (\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{vv}) & (\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{v}) \\ (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{vv}) & (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{v}) \\ (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{vv}) & (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{v}) \end{vmatrix} - \right.$$
$$\left. - \begin{vmatrix} (\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_{uv}) & (\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_{v}) \\ (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{uv}) & (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{v}) \\ (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{uv}) & (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{v}) \end{vmatrix}\right\}.$$

Пользуясь тем, что алгебраические дополнения левых верхних элементов выписанных определителей совпадают, и тем, что $\boldsymbol{f}_u\cdot\boldsymbol{f}_u = E$, $\boldsymbol{f}_u\cdot\boldsymbol{f}_v = F$, $\boldsymbol{f}_v\cdot\boldsymbol{f}_v = G$, получим окончательно

$$K(u,\,v) = \frac{1}{(EG-F^2)^2}\left\{\begin{vmatrix} (\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{vv} - \boldsymbol{f}_{uv}^2) & (\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{v}) \\ (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{vv}) & E & F \\ (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{vv}) & F & G \end{vmatrix} - \right.$$
$$\left. - \begin{vmatrix} 0 & (\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_{u}) & (\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_{v}) \\ (\boldsymbol{f}_{u}\cdot\boldsymbol{f}_{uv}) & E & F \\ (\boldsymbol{f}_{v}\cdot\boldsymbol{f}_{uv}) & F & G \end{vmatrix}\right\}.$$

Для того чтобы получить выражение гауссовой кривизны, включающее только функции E, F, G и их частные производные, достаточно выразить функции $\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_u$, $\boldsymbol{f}_u\cdot\boldsymbol{f}_{uv}$, $\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_v$, $\boldsymbol{f}_v\cdot\boldsymbol{f}_{vv}$, $\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_v$, $\boldsymbol{f}_u\cdot\boldsymbol{f}_{vv}$ и $\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{vv} - \boldsymbol{f}_{uv}^2$ через частные производные функций E, F, G. Начнем с того, что, продифференцировав по u и по v соотношения

$$\boldsymbol{f}_u^2 = E, \quad \boldsymbol{f}_v^2 = G, \quad \boldsymbol{f}_u\cdot\boldsymbol{f}_v = F,$$

получим следующие тождества:

$$\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_u = \frac{1}{2}E_u, \quad \boldsymbol{f}_u\cdot\boldsymbol{f}_{uv} = \frac{1}{2}E_v,$$

$$\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_v=\frac{1}{2}G_u,\quad \boldsymbol{f}_v\cdot\boldsymbol{f}_{vv}=\frac{1}{2}G_v,$$
$$\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_v+\boldsymbol{f}_u\cdot\boldsymbol{f}_{uv}=F_u,\quad \boldsymbol{f}_u\cdot\boldsymbol{f}_{vv}+\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_v=F_v.$$

Вычитая из последних двух равенств соответственно второе и третье равенства, получим:

$$\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_v=F_u-\frac{1}{2}E_v,\quad \boldsymbol{f}_u\cdot\boldsymbol{f}_{vv}=F_v-\frac{1}{2}G_u.$$

Далее, поскольку

$$(\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_v)_v=\boldsymbol{f}_{uuv}\cdot\boldsymbol{f}_v+\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{vv},\ (\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_v)_u=\boldsymbol{f}_{uuv}\cdot\boldsymbol{f}_v+\boldsymbol{f}_{uv}^2,$$

то

$$\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_{vv}-\boldsymbol{f}_{uv}^2=(\boldsymbol{f}_{uu}\cdot\boldsymbol{f}_v)_v-(\boldsymbol{f}_{uv}\cdot\boldsymbol{f}_v)_u=F_{uv}-\frac{1}{2}E_{vv}-\frac{1}{2}G_{uu}.$$

Подставляя все эти выражения в полученную нами ранее формулу для гауссовой кривизны, мы приходим наконец к формуле Гаусса:

$$K(u,\,v)=\frac{1}{(EG-F^2)^2}\times\\
\times\left\{\begin{vmatrix} F_{uv}-\frac{1}{2}(E_{vv}+G_{uu}) & \frac{1}{2}E_u & \left(F_u-\frac{1}{2}E_v\right)\\ F_v-\frac{1}{2}G_u & E & F\\ \frac{1}{2}G_u & F & G\end{vmatrix}-\right.\\
\left.-\begin{vmatrix} 0 & \frac{1}{2}E_v & \frac{1}{2}G_u\\ \frac{1}{2}E_v & E & F\\ \frac{1}{2}G_u & F & G\end{vmatrix}\right\}.$$

Основные уравнения теории поверхностей. Еще одно важное следствие из формулы Гаусса состоит в том, что коэффициенты первой и второй квадратичных форм поверхности не являются независимыми. Как мы видели, формула Гаусса позволяет выразить функцию $LN-M^2$ через функции E, F, G и их частные производные. Напрашивается вопрос: существуют ли еще какие-нибудь зависимости между коэффициентами первой и второй квадратичных форм? Ответ на этот вопрос положительный. Имеются две формулы Петерсона — Майнарди — Кодацци, выражающие линейные комбинации частных производных функций L, M, N через сами эти функции и через функции E, F, G и их частные производные:

$$L_v-M_u=\varphi(L,\,M,\,N;\ E,\,\ldots,\,G_v),$$

$$M_v - N_u = \psi(L,\, M,\, N;\; E,\, \ldots,\, G_v).$$

Других зависимостей между E, F, G, L, M, N *нет.*

Оказывается, что если в плоской области V заданы шесть функций L, M, N, E, F, G, для которых выполняются уравнения Петерсона — Майнарди — Кодацци, функция $LN - M^2$ удовлетворяет уравнению Гаусса и функции E, $EG - F^2$ всюду положительны, то указанные шесть функций являются коэффициентами первой и второй квадратичных форм некоторой гладкой поверхности, определенной ими однозначно с точностью до положения в пространстве. Это теорема Боннэ.

§ 14. Геодезическая кривизна и геодезические кривые

Определение. Пусть C — кривая на гладкой поверхности Φ, а $\boldsymbol{k}$ — вектор кривизны в точке P этой кривой. Вектор $\boldsymbol{k}$ раскладывается в сумму двух векторов $\boldsymbol{k}_g$ и $\boldsymbol{k}_n$, первый из которых лежит в касательной плоскости $T_P\Phi$, а второй ей перпендикулярен. Как мы знаем, длина вектора $\boldsymbol{k}_n$ равна абсолютной величине нормальной кривизны кривой C в точке P:

$$|\boldsymbol{k}_n| = |k_n|.$$

Длина вектора $\boldsymbol{k}_g$ называется *геодезической кривизной* (*без знака*) кривой C в точке P и обозначается через k_g:

$$|\boldsymbol{k}_g| = k_g.$$

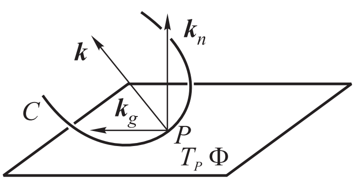


Рис. 89

В силу перпендикулярности векторов $\boldsymbol{k}_g$ и $\boldsymbol{k}_n$ имеем:

$$k^2 = k_g^2 + k_n^2,$$

где $k = |\boldsymbol{k}|$ — кривизна кривой C в точке P (рис. 89).

В качестве упражнения предлагаем читателю доказать, что геодезическая кривизна кривой C в точке P равна кривизне проекции этой кривой на касательную плоскость $T_P\Phi$. В частности, на плоскости геодезическая кривизна всякой кривой в каждой точке совпадает с ее кривизной (рис. 90).

Если кривая C параметризована вектор-функцией $\boldsymbol{\varphi}(t)$, причем $P = \boldsymbol{\varphi}(t_0)$, то, как нетрудно видеть, ее геодезическая кривизна в точке P может быть вычислена по формуле

$$k_g = \frac{|(\boldsymbol{\varphi}''(t_0),\, \boldsymbol{\varphi}'(t_0),\, \boldsymbol{n})|}{|\boldsymbol{\varphi}'(t_0)|^3},$$

где $\boldsymbol{n}$ — вектор нормали к поверхности в точке P.

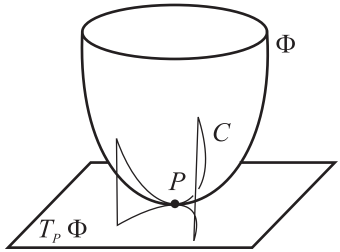


Рис. 90

Геодезические кривые. Оказывается, что геодезическая кривизна относится к внутренней геометрии. (Мы оставим этот факт без доказательства.) Ввиду этого особый интерес представляют кривые на поверхности, геодезическая кривизна которых в каждой точке равна нулю. Такие кривые называются ***геодезическими*** кривыми или просто — ***геодезическими***. Они являются аналогами отрезков прямых на плоскости, кривизна которых в каждой точке равна нулю.

Из-за важности геодезических для внутренней геометрии укажем несколько их свойств, каждое из которых может быть взято за определение. Если кривая C является геодезической, то:

а) вектор главной нормали кривой C в каждой точке совпадает с вектором нормали к поверхности (с точностью до знака) (рис. 91);

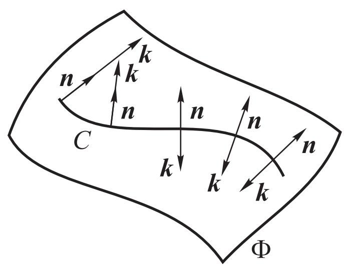


Рис. 91

б) соприкасающаяся плоскость в каждой точке кривой C проходит через нормаль к поверхности;

в) спрямляющая плоскость кривой C в каждой точке совпадает с касательной плоскостью к поверхности;

г) кривая C в каждой точке имеет наименьшую кривизну среди всех кривых, проходящих через эту же точку в том же направлении (т. е. она является «наипрямейшей» среди этих кривых);

д) если $\boldsymbol{\varphi}(t)$ — параметризация кривой C, то $(\boldsymbol{\varphi}''(t),\, \boldsymbol{\varphi}'(t),\, \boldsymbol{n}) \equiv 0$.

Как известно, через каждые две точки на плоскости проходит единственная прямая. Для геодезических на поверхности это свойство, вообще говоря, может не выполняться. Две точки на поверхности могут соединяться двумя, тремя, и даже бесконечным числом геодезических,

а могут и не соединяться ни одной. Однако выполнено другое, родственное свойство, а именно:

Теорема. *Через каждую точку поверхности в каждом направлении проходит единственная геодезическая.* (Поскольку любая дуга геодезической сама есть геодезическая, то единственность требует пояснений. Под единственностью мы имеем в виду, что любые две достаточно короткие геодезические, проходящие через одну и ту же точку в одном и том же направлении, являются дугами одной и той же большей геодезической; рис. 92.)

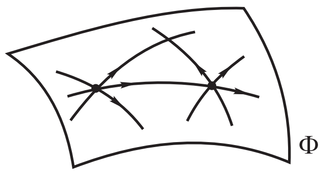


Рис. 92

Чтобы доказать это, составим дифференциальное уравнение геодезической. Пусть $\boldsymbol{f}\,(u,\ v)$ — параметризация поверхности Φ. Будем искать внутренние уравнения геодезической в виде $u = t$, $v = \psi(t)$. Тогда параметризация геодезической будет иметь вид

$$\boldsymbol{r} = \boldsymbol{\varphi}(t) = \boldsymbol{f}\,(t,\ \psi(t)).$$

Далее, по обычным формулам,

$$\begin{aligned}\boldsymbol{\varphi}'(t) &= \boldsymbol{f}_u(t,\ \psi(t)) + \boldsymbol{f}_v(t,\ \psi(t)) \cdot \psi'(t),\\ \boldsymbol{\varphi}''(t) &= \boldsymbol{f}_{uu}(t,\ \psi(t)) + 2\boldsymbol{f}_{uv}(t,\ \psi(t))\psi'(t) + \\ &\qquad + \boldsymbol{f}_{vv}(t,\ \psi(t))\psi'^2(t) + \boldsymbol{f}_v(t,\ \psi(t))\psi''(t).\end{aligned}$$

Для краткости опуская аргументы t и $\psi(t)$ у функций, можно записать:

$$\begin{aligned}(\boldsymbol{\varphi}'',\ \boldsymbol{\varphi}',\ \boldsymbol{n}) &= (\boldsymbol{f}_{uu} + 2\boldsymbol{f}_{uv}\psi' + \boldsymbol{f}_{vv}\psi'^2 + \boldsymbol{f}_v\psi'',\ \boldsymbol{f}_u + \boldsymbol{f}_v\psi',\ \boldsymbol{n}) = \\ &= (\boldsymbol{f}_{uu} + 2\boldsymbol{f}_{uv}\psi' + \boldsymbol{f}_{vv}\psi'^2,\ \boldsymbol{f}_u + \boldsymbol{f}_v\psi',\ \boldsymbol{n}) + (\boldsymbol{f}_v\psi'',\ \boldsymbol{f}_u + \boldsymbol{f}_v\psi',\ \boldsymbol{n}).\end{aligned}$$

Последнее слагаемое равняется $-\psi'' \cdot (\boldsymbol{f}_u,\ \boldsymbol{f}_v,\ \boldsymbol{n})$. Таким образом, равенство $(\boldsymbol{\varphi}'',\ \boldsymbol{\varphi}',\ \boldsymbol{n}) = 0$ равносильно следующему равенству:

$$\psi'' = \frac{(\boldsymbol{f}_{uu} + 2\boldsymbol{f}_{uv}\psi' + \boldsymbol{f}_{vv}\psi'^2,\ \boldsymbol{f}_u + \boldsymbol{f}_v\psi',\ \boldsymbol{n})}{(\boldsymbol{f}_u,\ \boldsymbol{f}_v,\ \boldsymbol{n})}.$$

Это равенство можно рассматривать как обыкновенное дифференциальное уравнение второго порядка относительно функции $\psi(t)$. Теперь из теоремы существования и единственности решений обыкновенных

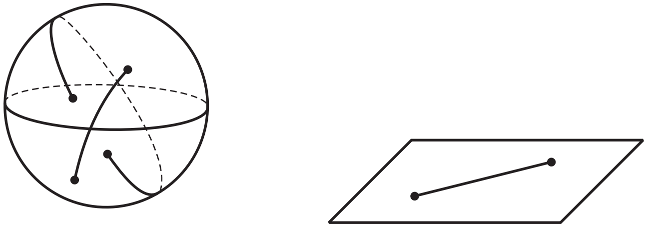


Рис. 93

дифференциальных уравнений следует, что через каждую точку в любом направлении проходит единственная геодезическая (если интересующее нас направление совпадает с направлением вектора $\boldsymbol{f}_v$, то мы можем поменять ролями координаты u и v). □

Прямые на плоскости, очевидно, являются геодезическими. Нетрудно убедиться, что на цилиндре геодезическими являются окружности, винтовые линии и прямые, а на сфере геодезическими являются окружности больших кругов. Поскольку через каждую точку этих поверхностей в каждом направлении проходит линия одного из указанных типов, то *других геодезических на плоскости, цилиндре и сфере нет* (рис. 93).

§ 15. Полугеодезическая параметризация поверхности. Экстремальное свойство геодезических

Определение. Параметризация $\boldsymbol{f}(u, v)$ гладкой поверхности Φ называется *полугеодезической*, если все координатные линии одного семейства являются геодезическими и в каждой точке поверхности координатные линии ортогональны (рис. 94).

Способ построения полугеодезической параметризации усматривается из определения. Пусть L — произвольная гладкая кривая на поверхности Φ, а $\boldsymbol{\varphi}(t)$ — ее регулярная (дважды дифференцируемая) параметризация. Через каждую точку $P = \boldsymbol{\varphi}(t)$ кривой L проходит вполне определенная геодезическая C_t, ортогональная кривой L (рис. 95). Будем считать ее ориентированной таким образом, что в точке P касательные векторы кривых C_t и L образуют с вектором нормали к Φ правую тройку. Введем на кривой C_t естественную параметризацию $\boldsymbol{\psi}_t$, в которой $\boldsymbol{\psi}_{(t)}(0) = \boldsymbol{\varphi}(t)$. Теперь мы можем рассмотреть вектор-функцию $\boldsymbol{g}(s, t)$, определенную в некоторой окрестности W от-

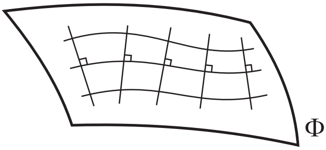


Рис. 94

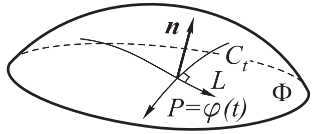


Рис. 95

резка $[a,\, b]$ в плоскости с координатами s, t по формуле

$$\boldsymbol{g}(s,\, t) = \boldsymbol{\psi}_{(t)}(s).$$

Из теоремы о зависимости решений дифференциального уравнения от коэффициентов и начальных данных следует, что $\boldsymbol{g}(s,\, t)$ — гладкая вектор-функция, а из теоремы об обратном отображении следует, что $\boldsymbol{g}\colon W \to \mathbf{R}^3$ есть регулярная параметризация некоторой окрестности Ω кривой L в поверхности Φ. Тем самым мы построили в окрестности кривой L систему внутренних координат s, t, в которой сама кривая L будет задаваться уравнением $s = 0$, а геодезическая C_{t_0} — уравнением $t = t_0$ (рис. 96).

Рис. 96

Чтобы убедиться в том, что $\boldsymbol{g}(s,\, t)$ есть полугеодезическая параметризация, осталось проверить ортогональность координатных линий в каждой точке, т. е. равенство $F \equiv 0$. Так как любая кривая C_{t_0} — геодезическая, то вектор $\boldsymbol{\psi}''_{(t_0)}(s)$ всюду направлен по нормали к поверхности. Но, очевидно, $\boldsymbol{g}_{ss}(s, t_0) = \boldsymbol{\psi}''_{(t_0)}(s)$. Поэтому $\boldsymbol{g}_{ss} \cdot \boldsymbol{g}_t \equiv 0$. Как мы знаем (см. § 13), $\boldsymbol{g}_{ss} \cdot \boldsymbol{g}_t = F_s - E_t/2$. Но при любом t_0

$$E(s,\, t_0) = \boldsymbol{g}_s^2(s,\, t_0) = [\boldsymbol{\psi}'_{(t_0)}(s)]^2 \equiv 1$$

в силу естественности параметризации $\boldsymbol{\psi}_{(t_0)}$; поэтому $E_t \equiv 0$ и $F_s = \boldsymbol{g}_{ss} \cdot \boldsymbol{g}_t + E_t/2 \equiv 0$. С другой стороны, по построению, $F(0,\, t) \equiv 0$.

Поэтому $F(s, t) \equiv 0$, и, следовательно, построенная параметризация $\boldsymbol{g}(s, t)$ действительно полугеодезическая. $\square$

Еще раз отметим, что в построенной параметризации $E \equiv 1$, $F \equiv 0$. Как нетрудно убедиться, всякая параметризация, удовлетворяющая этому условию, заведомо является полугеодезической.

Теорема. *Дуга геодезической кривой C между точками A и B будет кратчайшей среди всех кривых, лежащих на поверхности, с концами в этих точках, если точки A и B достаточно близки.*

Доказательство. Применим описанную выше конструкцию к случаю, когда кривая L ортогональна геодезической C в точке A. Мы получим в (возможно, очень маленькой) окрестности Ω точки A полугеодезическую параметризацию, которой соответствует система внутренних координат s, t. В этой системе кривая L будет задаваться уравнением $s = 0$. Если точка A имеет координаты $(0, t_0)$, то кривая C будет задаваться уравнением $t = t_0$. Пусть точка B имеет координаты $(0, s_0)$ (рис. 97).

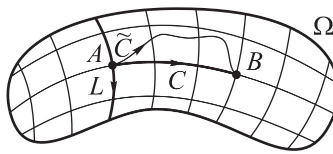


Рис. 97

Рассмотрим любую другую кривую $\widetilde{C}$ с концами в A и B. Для простоты предположим, что $\widetilde{C}$ — гладкая кривая, целиком лежащая в окрестности Ω. Пусть $u = \varphi_1(\tau)$, $v = \varphi_2(\tau)$ — ее внутренние уравнения, где $\tau \in [\alpha, \beta]$. При этом, очевидно, $\varphi_1(\alpha) = 0$, $\varphi_1(\beta) = s_0$. Если $S(\widetilde{C})$ и $S(C)$ — длины кривых $\widetilde{C}$ и C, то мы получаем неравенство:

$$S(\widetilde{C}) = \int\limits_{\alpha}^{\beta} \sqrt{[\varphi_1'(\tau)]^2 + G \cdot [\varphi_2'(\tau)]^2}\, d\tau \geqslant \int\limits_{\alpha}^{\beta} |\varphi_1'(\tau)|\, d\tau \geqslant$$

$$\geqslant \left| \int\limits_{\alpha}^{\beta} \varphi_1'(\tau)\, d\tau \right| = |\varphi_1(\beta) - \varphi_1(\alpha)| = |s_0| = S(C).$$

Ясно, что $S(\widetilde{C}) = S(C)$ тогда и только тогда, когда $\varphi_2(\tau) \equiv 0$, $\varphi_1'(\tau) > 0$, т. е. $\widetilde{C} = C$. Теорема доказана. $\square$

Заключительные замечания: поверхности постоянной гауссовой кривизны. Пользуясь полугеодезической параметризацией и формулой Гаусса, можно доказать, что если у двух поверхностей в каждой точке гауссова кривизна равна одному и тому же числу K,

то эти поверхности *локально изометричны*: любой достаточно малый участок первой поверхности изометричен некоторому участку второй, и наоборот. Поэтому на поверхностях постоянной гауссовой кривизны $K \geqslant 0$ «в малом» выполняется сферическая геометрия (при $K > 0$) или планиметрия (при $K = 0$). При $K < 0$ на поверхности реализуется «в малом» геометрия Лобачевского, подробно излагаемая в последней главе учебника. С виду эти три геометрии совершенно различны, что проявляется, например, в поведении суммы углов треугольника. Если на евклидовой плоскости она равна π, то на сфере — всегда больше, а на плоскости Лобачевского — всегда меньше π, причем в обоих случаях отличие суммы углов от π прямо пропорционально *площади* рассматриваемого треугольника. Общее объяснение этому дает теорема Гаусса—Боннэ. Она утверждает, что если α, β, γ — внутренние углы треугольной области Δ, ограниченной на гладкой поверхности тремя геодезическими, то

$$\alpha + \beta + \gamma = \pi + \iint\limits_{\Delta} K\,dS.$$

(В этой формуле интеграл понимается как «интеграл по площади», который можно определить формально как $\iint\limits_{\widetilde{D}} K(u, v)\sqrt{EG - F^2}du\,dv$, где $\widetilde{D}$ — часть области параметров, соответствующая области Δ.) Для поверхности постоянной гауссовой кривизны второе слагаемое в правой части очевидно равно произведению гауссовой кривизны K на площадь треугольника Δ. Приведенную формулу легко обобщить на случай геодезического многоугольника.

Часть 5
ТОПОЛОГИЯ

С первыми топологическими понятиями мы уже встречались во второй части книги, где определялись внутренние и граничные точки множества и т. п. При общем изучении топологии за основу, оказывается, удобнее всего взять понятие открытого множества.

Здесь мы начинаем с того, что определяем топологическое пространство и изучаем его «геометрию» исходя из аксиом. После того, как введены гомеоморфизмы, изучаются важнейшие топологические свойства пространств и множеств — связность и компактность. В заключение описывается классификация связных компактных двумерных многообразий, к которым относятся такие интересные объекты, как лист Мебиуса, бутылка Клейна, проективная плоскость, тор, крендель и сферы с ручками.

Глава I
ТОПОЛОГИЧЕСКИЕ ПРОСТРАНСТВА И НЕПРЕРЫВНЫЕ ОТОБРАЖЕНИЯ

§ 1. Топология в множестве

Пусть X — произвольное множество. *Топологической структурой*, или *топологией*, в множестве X называется совокупность Ω его подмножеств, для которой выполнены три условия:

а) объединение любого семейства множеств, принадлежащих совокупности Ω, также принадлежит совокупности Ω;

б) пересечение любых двух множеств, принадлежащих совокупности Ω, также принадлежит совокупности Ω;

в) пустое множество $\varnothing$ и все множество X принадлежат совокупности Ω.

Множество X с выделенной топологической структурой Ω называется *топологическим пространством* и обозначается (X, Ω) или просто X, если ясно, о какой топологической структуре идет речь. Элементы множества X называются *точками* пространства (X, Ω). Множества, входящие в выделенную совокупность Ω, называются *открытыми* в X множествами[15]. Условия а) — в) называются *аксиомами топологической структуры.* Они выражают следующие основные свойства открытых множеств:

а) *объединение любого семейства открытых множеств открыто*;

б) *пересечение любых двух открытых множеств открыто*;

в) *пустое множество и все пространство открыты.*

Заметим, что, как следует из б) по индукции,

б') *пересечение любого конечного числа открытых множеств открыто.*

Запишем аксиомы а) — в) еще и при помощи формул.

а) Пусть $\{U_\alpha\}_{\alpha\in I}$ — семейство множеств U_α, где индекс α пробегает множество индексов I. Если $U_\alpha \in \Omega$ для каждого индекса α из I, то и $\bigcup\limits_{\alpha\in I} U_\alpha \in \Omega$.

б) Если $U_1, U_2 \in \Omega$, то и $U_1 \cap U_2 \in \Omega$.

в) $\varnothing, X \in \Omega$.

Наконец, утверждение б') означает, что если $U_1, U_2, \ldots, U_n \in \Omega$, то и $\bigcap\limits_{i=1}^{n} U_i \in \Omega$.

Рассмотрим простейшие примеры топологических пространств.

1. Если Ω совпадает с множеством всех подмножеств множества X, то топологическое пространство (X, Ω) называется *дискретным.* Мы видим, что в дискретном пространстве все множества открыты.

2. Если Ω содержит всего два множества: $\varnothing$ и X, то топологическое пространство (X, Ω) называется *антидискретным* пространством или пространством с *тривиальной топологией.* В антидискретном пространстве только два открытых множества: пустое множество и все пространство. (Дискретное пространство можно сравнить с мешком гороха, где каждая горошинка — сама по себе, а антидискретное — с запутанным клубком ниток или тарелкой спагетти. Здесь роль точек играют отдельные горошины и макаронины.)

---

[15] Буква Ω соответствует букве O, с которой начинаются слова: открытый, open (англ.), offen (нем.), ouvert (франц.), все означающие одно и то же.

3. Пусть X есть луч $[0, +\infty)$, а Ω состоит из $\varnothing$, X и всевозможных лучей $(a, +\infty)$, где $a \geqslant 0$.

Для совокупности Ω аксиомы б) и в) очевидно выполнены, а аксиома а) просто означает, что объединение любого семейства таких лучей снова есть луч. Пространство (X, Ω) называется *стрелкой*.

Дискретная и антидискретная топологии сами по себе не слишком интересны, но мы постоянно будем пользоваться ими в качестве простейших примеров.

Дополнения открытых множеств имеют специальное название.

Множество $F \subset X$ называется *замкнутым* в пространстве (X, Ω), если его дополнение $X \setminus F$ открыто, т. е. $X \setminus F \in \Omega$.

Примеры. 1. В дискретном пространстве все множества замкнуты.

2. В антидискретном пространстве только два замкнутых множества: пустое множество и все пространство.

3. В стрелке замкнуты пустое множество $\varnothing$, весь луч $[0, +\infty)$ и отрезки вида $[0, a]$, где $a \geqslant 0$.

Перечислим свойства замкнутых множеств, вытекающие из аксиом топологической структуры.

Теорема.

а) *Пересечение любого набора замкнутых множеств замкнуто.*

б) *Объединение любых двух замкнутых множеств замкнуто.*

в) *Пустое множество и все пространство замкнуты.*

Из утверждения б) следует по индукции, что

б') *Объединение любого конечного числа замкнутых множеств замкнуто.*

Доказательство. Начнем со свойства в). Оно вытекает из аксиомы в) топологической структуры: пустое множество $\varnothing$ замкнуто, поскольку открыто его дополнение — все пространство X; а все пространство X замкнуто, поскольку открыто его дополнение — пустое множество $\varnothing$.

Перейдем к свойству б). Пусть F, $G \subset X$ — замкнутые множества. Это значит, что их дополнения $X \setminus F$ и $X \setminus G$ открыты. Чтобы доказать, что объединение $F \cup G$ тоже замкнуто, нужно проверить, что дополнение множества $F \cup G$ открыто. Но дополнение объединения двух множеств совпадает с пересечением их дополнений:

$$X \setminus (F \cup G) = (X \setminus F) \cap (X \setminus G).^{16}$$

[16] Для обозначения дополнения множества иногда используется значок $\complement$ (стили-

Поэтому множество $X \setminus (F \cup G)$ открыто как пересечение двух открытых множеств (аксиома б) топологической структуры).

Докажем, наконец, свойство а). Пусть $\{F_\alpha\}_{\alpha \in I}$ — произвольное семейство замкнутых множеств пространства. Это значит, что открыты их дополнения — множества $X \setminus F_\alpha$, $\alpha \in I$. Чтобы доказать, что пересечение $\bigcap\limits_{\alpha \in I} F_\alpha$ тоже замкнуто, нужно проверить, что дополнение множества $\bigcap\limits_{\alpha \in I} F_\alpha$ открыто. Но дополнение пересечения нескольких множеств совпадает с объединением их дополнений:

$$X \setminus \bigcap_{\alpha \in I} F_\alpha = \bigcup_{\alpha \in I} (X \setminus F_\alpha).$$

Поэтому множество $X \setminus \bigcap\limits_{\alpha \in I} F_\alpha$ открыто как объединение нескольких открытых множеств (аксиома а) топологической структуры). $\square$

На примере дискретного и антидискретного пространств видно, что множество может быть одновременно открытым и замкнутым, а может не быть ни замкнутым, ни открытым (как закрытая, но не запертая дверь!).

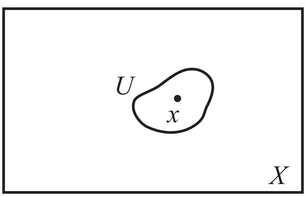


Рис. 1

Свойства замкнутых и открытых множеств в чем-то похожи. Важное различие между ними в том, что пересечение бесконечного семейства открытых множеств не обязательно открыто, а пересечение бесконечного семейства замкнутых множеств всегда замкнуто; в то же время объединение бесконечного семейства замкнутых множеств не обязательно замкнуто, а объединение бесконечного семейства открытых множеств всегда открыто. Например, на числовой прямой $\mathbf{R}^1$ имеем

$$\bigcap_{n=1}^{\infty} \left(-1 - \frac{1}{n},\, 1 + \frac{1}{n} \right) = [-1,\, 1],$$

$$\bigcup_{n=1}^{\infty} \left[-1 + \frac{1}{n},\, 1 - \frac{1}{n} \right] = (-1,\, 1)!$$

зованная первая буква слова complement — дополнение). Тогда запись становится симметричной: $\complement(F \cup G) = \complement F \cap \complement G$.

Наконец, дадим еще одно определение. *Топологической окрестностью точки* x топологического пространства X называется всякое открытое множество U, содержащее эту точку x (рис. 1). Часто мы будем опускать слово «топологическая» и говорить просто «окрестность».

§ 2. Метрика в множестве

Пусть M — произвольное множество. *Метрикой* в множестве M называется вещественная функция ρ, определенная на множестве всевозможных пар элементов множества M:

$$\rho : M \times M \to \mathbf{R}, \quad (x,\ y) \longmapsto \rho(x,\ y),$$

такая, что выполнены четыре условия:

а) функция ρ принимает только неотрицательные значения: $\rho(x, y) \geqslant 0$ для любых x, y из M;

б) $\rho(x, x) = 0$ для любого элемента x из M, и если $\rho(x, y) = 0$, то обязательно $x = y$;

в) $\rho(x, y) = \rho(y, x)$ для любых x, y из M;

г) $\rho(x, z) \leqslant \rho(x, y) + \rho(y, z)$ для любых x, y, z из M.

Множество M с фиксированной метрикой ρ называется *метрическим пространством* и обозначается (M, ρ) или просто M, если ясно, о какой метрике идет речь. Элементы множества M называются *точками* пространства (M, ρ). Значение метрической функции ρ на паре элементов x, y называется *расстоянием* между точками x и y и обозначается обычно символом $\mathrm{dist}(x, y)$: $\mathrm{dist}(x, y) = \rho(x, y)$.[17] Это обозначение позволяет не указывать метрическую функцию явно. Условия а) — г) называются *аксиомами метрики.* Они выражают основные свойства расстояния:

а) Неотрицательность: *расстояние между двумя точками всегда неотрицательно.*

б) Аксиома тождества: *расстояние между точками равно нулю тогда и только тогда, когда точки совпадают.*

в) Симметричность: *расстояние от точки x до точки y равно расстоянию от точки y до точки x.*

Условие г) называется *неравенством треугольника*, поскольку оно аналогично тому факту, что длина любой стороны треугольника меньше суммы длин двух других его сторон.

---

[17] От distance (фр., англ.). Ср. русское «дистанция».

Рассмотрим примеры метрических пространств.

1. Возьмем в качестве множества M произвольное множество и для любых x, y из M положим $\rho(x, y) = 1$, если $x \neq y$, и $\rho(x, y) = 0$, если $x = y$. Такая метрика называется *симплициальной.* В пространстве с симплициальной метрикой расстояние между любыми двумя различными точками равно 1.

2. Возьмем в качестве множества M множество всех вещественных чисел $\mathbf{R}$. Определим метрику ρ по формуле

$$\rho(x, y) = |x - y|.$$

Это *стандартная метрика* на прямой.

3. Возьмем в качестве M n-мерное евклидово пространство $\mathbf{R}^n$, а в качестве метрики ρ — обычное *евклидово расстояние* между точками: если $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ — точки пространства $\mathbf{R}^n$, то

$$\rho(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}.$$

4. Рассмотрим в качестве M n-мерное проективное пространство, т. е. множество всех прямых в $(n+1)$-мерном евклидовом пространстве $\mathbf{R}^{n+1}$, проходящих через начало координат. При $n = 2$ это проективная плоскость, а при $n = 3$ — обычное проективное пространство. Расстояние между двумя «точками» — прямыми l_1 и l_2 положим равным углу между l_1 и l_2. Первые три аксиомы метрики здесь очевидно выполнены. Четвертая аксиома — неравенство треугольника — выражает тот факт, что в трехгранном угле каждый плоский угол меньше суммы двух остальных (см. ч. 2, гл. III, § 4, теорема 1).

Евклидово n-мерное пространство — в первую очередь плоскость и обычное трехмерное пространство — будут для нас основными примерами метрических пространств. Теорию метрических пространств можно считать разновидностью геометрии, построенной исключительно на понятии расстояния.

Всякое подмножество A метрического пространства M можно рассматривать как самостоятельное метрическое пространство. Для этого нужно определить расстояние между каждыми двумя его точками как расстояние между этими же точками в исходном пространстве M. Такая метрика называется *индуцированной.* Подмножество A, наделенное индуцированной метрикой, называется *метрическим подпространством* метрического пространства M.

Пусть M — метрическое пространство, a — его точка, r — положительное число. Множество точек пространства M, удаленных от точ-

ки a на расстояние, не большее r, называется *замкнутым* шаром пространства M с *центром* в точке a и *радиусом* r и обозначается символом $\mathrm{D}_r(a)$.[18] Множество точек, удаленных от точки a на расстояние, меньшее r, называется *открытым шаром* и обозначается $\mathrm{B}_r(a)$.[19] Наконец, множество точек, расположенных на расстоянии r от точки a, называется *сферой* и обозначается $\mathrm{S}_r(a)$.[20] Таким образом,

$$\mathrm{D}_r(a) = \{x \in M\colon \operatorname{dist}(a, x) \leqslant r\},$$
$$\mathrm{B}_r(a) = \{x \in M\colon \operatorname{dist}(a, x) < r\},$$
$$\mathrm{S}_r(a) = \{x \in M\colon \operatorname{dist}(a, x) = r\}.$$

Единичные замкнутый шар и сферу в евклидовом пространстве $\mathbf{R}^n$ обозначим через D^n и S^{n-1}. В случае когда M — обычное евклидово пространство, метрические сферы и замкнутые шары представляют собой хорошо знакомые нам геометрические фигуры — обычные сферы и шары.

Открытый шар радиуса $\varepsilon > 0$ с центром в данной точке часто также называется *ε-окрестностью* этой точки.

Подмножество A метрического пространства M называют *ограниченным*, если всевозможные попарные расстояния между его точками не превосходят некоторого фиксированного числа d. Ясно, что в этом случае все множество A содержится в шаре радиуса d с центром в произвольной точке множества A.[21] Наоборот, если множество A содержится в некотором шаре радиуса r с центром в некоторой точке a, то оно ограничено в силу неравенства треугольника. Действительно, пусть x, y — две точки из A. Тогда

$$\operatorname{dist}(x,\ y) \leqslant \operatorname{dist}(x,\ a) + \operatorname{dist}(a,\ y) \leqslant 2r.$$

Поэтому в качестве d можно взять число $2r$.

Опишем теперь некоторую выделенную топологическую структуру, которая существует во всяком метрическом пространстве M. Пусть $\Omega(M)$ — семейство всех множеств U, которые вместе с каждой своей точкой x содержат некоторую ее шаровую окрестность, т. е.

$$U \in \Omega(M) \Leftrightarrow \forall x \in U\ \exists r > 0\colon\ \mathrm{B}_r(x) \subset U.$$

---

[18] От disc — диск.

[19] От Ball (нем.), ball (англ.), balle (фр.).

[20] От Sphäre (нем.), sphere (англ.), sphère (фр.).

[21] Разумеется, если множество A непусто. Пустое множество считается ограниченным по определению.

Пустое множество тоже считается принадлежащим семейству $\Omega(M)$. (Пустое множество $\varnothing$ никаких точек не содержит, и потому, как принято считать, «все его точки» обладают любым наперед заданным свойством. В частности, они содержатся в пустом множестве вместе со всеми своими шаровыми окрестностями.)

Теорема 1. *Для совокупности $\Omega(M)$ выполнены аксиомы топологической структуры.*

Доказательство. Начнем с аксиомы а). Пусть $\{U_\alpha\}_{\alpha\in I}$ — произвольное семейство множеств, принадлежащих совокупности $\Omega(M)$. Чтобы доказать, что их объединение $\bigcup\limits_{\alpha\in I} U_\alpha$ также принадлежит совокупности $\Omega(M)$, найдем для произвольной точки $a \in \bigcup\limits_{\alpha\in I} U_\alpha$ некоторую ее шаровую окрестность V, содержащуюся в множестве $\bigcup\limits_{\alpha\in I} U_\alpha$. Воспользуемся тем, что $a \in U'_\alpha$ для некоторого индекса α' из I. В множестве U'_α вместе с точкой a содержится некоторая ее шаровая окрестность V. Тем более, V содержится в множестве $\bigcup\limits_{\alpha\in I} U_\alpha$, и, значит, окрестность V искомая.

Перейдем к аксиоме б). Пусть U_1, U_2 — два множества, принадлежащие совокупности $\Omega(M)$. Чтобы доказать, что их пересечение $U_1 \cap U_2$ тоже принадлежит совокупности $\Omega(M)$, найдем для произвольной точки a из $U_1 \cap U_2$ некоторую ее шаровую окрестность, содержащуюся в множестве $U_1 \cap U_2$. Пусть множество U_1 содержит ε_1-окрестность точки a, а множество U_2 содержит ε_2-окрестность точки a, причем $\varepsilon_1, \varepsilon_2 > 0$.

Обозначим через ε меньшее из чисел ε_1, ε_2. Тогда, очевидно, ε-окрестность точки a содержится в каждом из множеств U_1 и U_2, а значит, и в их пересечении $U_1 \cap U_2$.

Наконец, проверим аксиому в). Все пространство M заведомо принадлежит совокупности Ω, поскольку с каждой своей точкой содержит все ее шаровые окрестности. Пустое множество $\varnothing$ также принадлежит совокупности $\Omega(M)$. Теорема 1 доказана. $\square$

Итак, мы проверили, что совокупность $\Omega(M)$ задает топологию в множестве M, которая называется *метрической топологией*. О метрической топологии часто говорят, как о топологии, которую *порождает* метрика. Множества, входящие в совокупность $\Omega(M)$, называются *открытыми подмножествами* метрического пространства M. Таким образом, открытыми в метрическом пространстве называются множе-

ства, которые вместе с каждой своей точкой содержат все достаточно близкие к ней точки.

Докажем, что *всякий открытый шар* $\mathrm{B}_r(a)$ *является открытым множеством.*

Пусть b — произвольная точка этого шара. Тогда $\operatorname{dist}(a, b) < r$. Положим $\varepsilon = r - \operatorname{dist}(a, b)$. Достаточно убедиться в том, что точка b содержится в шаре $\mathrm{B}_r(a)$ вместе со своей ε-окрестностью $\mathrm{B}_\varepsilon(b)$. Для этого проверим, что в шаре $\mathrm{B}_r(a)$ содержится каждая точка $x \in \mathrm{B}_\varepsilon(b)$. Но это следует из неравенства треугольника и того, что $\operatorname{dist}(x, b) < \varepsilon$:

$$\operatorname{dist}(x, a) \leqslant \operatorname{dist}(x, b) + \operatorname{dist}(b, a) < \varepsilon + \operatorname{dist}(b, a) = r.$$

В частности, *всякая ε-окрестность точки метрического пространства является ее топологической окрестностью* (*в метрической топологии*).

Пусть (X, Ω) — топологическое пространство. Если топологическая структура Ω порождается некоторой метрикой в множестве X, то топологическое пространство (X, Ω) называется *метризуемым.* Две метрики, порождающие одну и ту же топологию, называются *эквивалентными.*

§ 3. Внутренность, замыкание, граница

Пусть A — подмножество топологического пространства X.

Среди открытых множеств, содержащихся в множестве A, есть наибольшее: это объединение всех таких множеств. Оно называется *внутренней частью*, или *внутренностью*, множества A и обозначается символом $\operatorname{int} A$ или, подробнее, $\operatorname{int}_X A$.[22]

Например, в случае когда $X = \mathbf{R}$ — числовая прямая, имеем $\operatorname{int}[0, 1] = (0, 1)$.

Всякое открытое множество, очевидно, совпадает со своей внутренней частью, и наоборот, если множество совпадает со своей внутренностью, то оно открыто:

$$A = \operatorname{int} A \Leftrightarrow A \in \Omega.$$

Подобным образом среди всех замкнутых множеств, содержащих множество A, есть наименьшее: это пересечение всех таких множеств. Оно

[22] От interieur (фр.) и interior (англ.). Другое распространенное обозначение для внутренности — $\mathring{A}$.

называется *замыканием* множества A и обозначается символом $\operatorname{cl} A$ или, подробнее, $\operatorname{cl}_X A$.[23]

Например, когда $X = \mathbf{R}$ — числовая прямая, получаем $\operatorname{cl}(0, 1) = [0, 1]$.

Всякое замкнутое множество, очевидно, совпадает со своим замыканием, и наоборот: если множество совпадает со своим замыканием, то оно замкнуто:

$$A = \operatorname{cl} A \Leftrightarrow A \text{ замкнуто}.$$

Очевидно, что внутренность множества A содержится в его замыкании:

$$\operatorname{int} A \subset \operatorname{cl} A.$$

Точки, принадлежащие замыканию множества A, но не принадлежащие его внутренности, образуют *границу* множества A. Граница обозначается символом $\operatorname{fr} A$ или, подробнее, $\operatorname{fr}_X A$:[24]

$$\operatorname{fr} A = \operatorname{cl} A \setminus \operatorname{int} A.$$

Например, в случае, когда $X = \mathbf{R}$ — числовая прямая, имеем $\operatorname{fr}[0, 1] = \operatorname{fr}(0, 1) = \{0, 1\}$, т. е. граница интервала состоит из двух точек — его концов.

Так как разность двух множеств равна пересечению первого множества с дополнением второго множества, то

$$\operatorname{fr} A = \operatorname{cl} A \setminus \operatorname{int} A = \operatorname{cl} A \cap (X \setminus \operatorname{int} A).$$

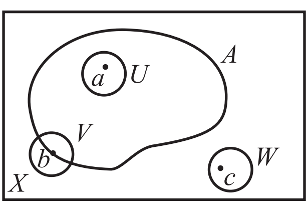


Рис. 2

Следовательно, граница $\operatorname{fr} A$ замкнута как пересечение двух замкнутых множеств.

Расположение точки относительно множества. Принадлежность произвольной точки пространства замыканию, внутренности или границе множества может быть описана на языке окрестностей.

Определение. Пусть $x \in X$ — произвольная точка. Точка x называется (рис. 2):

---

[23]От clôture (фр.) и closure (англ.). Другое распространенное обозначение для замыкания — $\overline{A}$.

[24]От frontière (фр.) и frontier (англ.). Другое распространенное обозначение для границы — ∂A.

а) *внутренней точкой* множества A, если она лежит в A вместе с некоторой своей окрестностью;

б) *граничной точкой* множества A, если всякая ее окрестность пересекается и с множеством A, и с его дополнением $X \setminus A$;

в) *внешней точкой* множества A, если она лежит в дополнении множества A вместе с некоторой своей окрестностью.

Ясно, что каждая точка $x \in X$ является по отношению к множеству A либо внутренней, либо граничной, либо внешней. Внутренние и граничные точки называются *точками прикосновения* множества A. Таким образом, точка $x \in X$ является точкой прикосновения множества A, если всякая ее окрестность пересекается с множеством A.

Теорема 1. а) *Внутренность* $\operatorname{int} A$ *всякого множества* A *есть множество его внутренних точек.*

б) *Замыкание* $\operatorname{cl} A$ *всякого множества* A *есть множество его точек прикосновения.*

в) *Граница* $\operatorname{fr} A$ *всякого множества* A *есть множество его граничных точек.*

Доказательство. а) Так как внутренность $\operatorname{int} A$ открыта, то она является окрестностью для всякой своей точки. Поэтому все ее точки — внутренние для множества A. Следовательно, множество $\operatorname{int} A$ содержится в множестве внутренних точек множества A.

Докажем обратное включение. Пусть $a \in A$ — произвольная внутренняя точка. Она содержится в A вместе с некоторой своей окрестностью U. Так как $\operatorname{int} A$ — наибольшее открытое множество, содержащееся в A, то окрестность U, а с ней и сама точка a содержится в множестве $\operatorname{int} A$. Следовательно, множество $\operatorname{int} A$ содержит множество внутренних точек множества A, а значит, в силу доказанного выше, и совпадает с ним.

б) Чтобы доказать, что множество $\operatorname{cl} A$ совпадает с множеством точек прикосновения множества A, проверим, что совпадают их дополнения. Дополнение к множеству точек прикосновения совпадает с множеством внутренних точек множества $X \setminus A$, т. е. с множеством $\operatorname{int}(X \setminus A)$. Но нетрудно показать, что и

$$X \setminus \operatorname{cl} A = \operatorname{int}(X \setminus A).$$

в) С очевидностью следует из а) и б) и того, что, во-первых, $\operatorname{fr} A = \operatorname{cl} A \setminus \operatorname{int} A$, и, во-вторых, граничные точки множества — это те точки прикосновения, которые не являются внутренними. Теорема 1 доказана. $\square$

§ 4. Подпространства топологического пространства

Подобно тому как всякое подмножество метрического пространства наделяется метрикой, так всякое подмножество топологического пространства может быть наделено естественной топологической структурой. (Сравните с тем, что далеко не всякое подмножество группы является ее подгруппой, подмножество линейного пространства — его подпространством.)

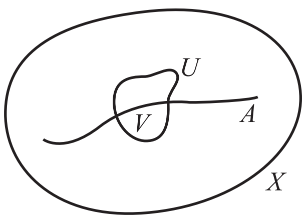


Рис. 3

Пусть A — произвольное подмножество топологического пространства (X, Ω) (рис. 3). Обозначим через Ω_A совокупность всех подмножеств множества A, которые являются пересечениями открытых множеств пространства X с множеством A:

$$\Omega_A = \{U \cap A \colon U \subset \Omega\}.$$

Теорема 1. *Совокупность Ω_A удовлетворяет аксиомам топологической структуры.*

Доказательство. Начнем с аксиомы а). Пусть $\{V_\alpha\}_{\alpha \in I}$ — семейство множеств, принадлежащих совокупности Ω_A. Это значит, что каждое множество V_α является пересечением некоторого открытого в X множества U_α с множеством A:

$$V_\alpha = U_\alpha \cap A.$$

Чтобы доказать, что их объединение $\bigcup\limits_{\alpha \in I} V_\alpha$ тоже принадлежит совокупности Ω_A, представим множество $\bigcup\limits_{\alpha \in I} V_\alpha$ в виде пересечения некоторого открытого множества с множеством A. Действительно,

$$\bigcup_{\alpha \in I} V_\alpha = \bigcup_{\alpha \in I} (U_\alpha \cap A) = \left(\bigcup_{\alpha \in I} U_\alpha\right) \cap A,$$

где множество $\bigcup\limits_{\alpha \in I} U_\alpha$ открыто как объединение нескольких открытых множеств.

Перейдем к аксиоме б). Пусть V_1, V_2 — два множества из совокупности Ω_A. Чтобы доказать, что их пересечение $V_1 \cap V_2$ также принадлежит совокупности Ω_A, представим множество $V_1 \cap V_2$ в виде пересечения множества A с некоторым открытым множеством пространства

X. Пусть $V_1 = U_1 \cap A$ и $V_2 = U_2 \cap A$, где U_1, U_2 — открытые множества. Тогда

$$V_1 \cap V_2 = (U_1 \cap A) \cap (U_2 \cap A) = (U_1 \cap U_2) \cap A,$$

где множество $U_1 \cap U_2$ открыто как пересечение двух открытых множеств (в Ω_X).

Наконец, аксиома в) очевидна: множества A и $\varnothing$ тривиальным образам представляются в виде пересечения открытых множеств пространства X с множеством A:

$$A = X \cap A, \quad \varnothing = \varnothing \cap A. \ \square$$

Итак, мы проверили, что совокупность Ω_A задает топологию в множестве A. Говорят, что топология Ω_A *индуцируется* топологией Ω, и называют ее *индуцированной топологией*. Множество A, снабженное индуцированной топологией, называется *топологическим подпространством* топологического пространства (X, Ω) и обозначается (A, Ω_A), или просто A, если ясно, что, речь идет именно об индуцированной топологической структуре. Множества, входящие в совокупность Ω_A, называются *открытыми в множестве* A. Итак, открытыми в A являются множества, высекаемые в нем множествами, открытыми в пространстве X. Как обычно, множество $G \subset A$ называется *замкнутым в множестве* A, если его дополнение $A \setminus G$ открыто в A.

Теорема 2. *В множестве A замкнуты те и только те множества, которые являются пересечениями замкнутых множеств пространства X с множеством A.*

Доказательство. Пусть F — замкнутое множество пространства X. Чтобы доказать, что множество $F \cap A$ замкнуто в A, нужно проверить, что его дополнение $A \setminus (F \cap A) = A \setminus F$ открыто в множестве A. Для этого представим множество $A \setminus F$ в виде пересечения множества A с некоторым открытым множеством пространства X:

$$A \setminus F = A \cap (X \setminus F),$$

где множество $X \setminus F$ открыто как дополнение замкнутого множества F.

Наоборот, пусть множество $G \subset A$ замкнуто в A. Это значит, что его дополнение $A \setminus G$ открыто в A. Это, в свою очередь, означает, что $A \setminus G = A \cap U$ для некоторого множества U, открытого в X. Множество $X \setminus U$, замкнутое в пространстве X, и дает в пересечении

с множеством A множество G. Чтобы убедиться в этом, воспользуемся цепочкой равенств:

$$A \cap (X \setminus U) = A \setminus (A \cap U) = A \setminus (A \setminus G) = G. \ \square$$

§ 5. Непрерывные отображения

Непрерывность «в целом». Отображение одного топологического пространства в другое называется *непрерывным в целом* (или просто — *непрерывным*), если при этом отображении прообраз любого открытого множества открыт. Таким образом, если (X, Ω_X) и (Y, Ω_Y) — топологические пространства, а $f\colon X \to Y$ — отображение, то f непрерывно, если

$$U \in \Omega_Y \Rightarrow f^{-1}(U) \in \Omega_X.$$

Примеры. 1. *Тождественное отображение* любого топологического пространства X в себя непрерывно. Оно обозначается id_X:

$$\mathrm{id}_X\colon\ X \to X, \quad x \longmapsto x.$$

2. Постоянное отображение всегда непрерывно. Пусть X, Y — топологические пространства, $y_0 \in Y$ — точка, а f — постоянное отображение:

$$f\colon\ X \to Y, \quad x \longmapsto y_0.$$

Тогда прообраз любого открытого множества $U \subset Y$ совпадает с X, если $y_0 \in U$, и пуст в противном случае.

3. Любое отображение дискретного пространства в любое топологическое пространство непрерывно.

4. Любое отображение любого топологического пространства в антидискретное пространство непрерывно.

5. *Отображение включения* в топологическое пространство X любого его подпространства A непрерывно. Оно обозначается через in_A:

$$\mathrm{in}_A\colon\ A \to X, \quad a \longmapsto a.$$

При отображении включения прообраз любого множества $B \subset X$ совпадает с пересечением этого множества и подпространства A:

$$\mathrm{in}_A^{-1}(B) = B \cap A.$$

Поэтому прообраз любого открытого множества в X открыт в подпространстве A по определению индуцированной топологии в A. Это доказывает, что отображение in_A непрерывно.

Непрерывность отображения в целом можно определить и при помощи замкнутых множеств.

Теорема 1. *Отображение $f\colon X \to Y$ непрерывно в целом тогда и только тогда, когда при этом отображении прообраз любого замкнутого множества A замкнут.*

Доказательство. Пусть отображение f непрерывно. Чтобы доказать, что множество $f^{-1}(A)$ замкнуто, проверим, что его дополнение $X \setminus f^{-1}(A)$ открыто. Действительно, оно совпадает с прообразом дополнения множества A:

$$X \setminus f^{-1}(A) = f^{-1}(Y \setminus A).$$

А множество $f^{-1}(Y \setminus A)$ открыто как прообраз открытого множества $Y \setminus A$ при непрерывном отображении f.

В обратную сторону доказательство проводится аналогично. $\square$

Замечание. Важно отметить, что, в отличие от прообразов, при непрерывных в целом отображениях образы открытых (замкнутых) множеств вовсе не обязательно будут открытыми (замкнутыми). Чтобы убедиться в этом, вернитесь к примерам 3 и 4.

Отметим простое, но очень важное свойство непрерывных отображений.

Теорема 2. *Композиция непрерывных отображений непрерывна.*

Более подробно, если X, Y, Z — топологические пространства, а $f\colon X \to Y$ и $g\colon Y \to Z$ — непрерывные отображения, то их композиция $g \circ f\colon X \to Z$ есть непрерывное отображение.

Доказательство. Обозначим через h сквозное отображение: $h = g \circ f$. Для проверки его непрерывности требуется доказать, что если $U \subset Z$ — произвольное открытое множество, то его прообраз $h^{-1}(U)$ открыт в X. В самом деле, множество $g^{-1}(U)$ открыто в Y в силу непрерывности отображения g, а его прообраз $f^{-1}(g^{-1}(U))$ открыт в X в силу непрерывности отображения f. Осталось заметить, что $h^{-1}(U) = f^{-1}(g^{-1}(U))$.

Следствие. *Сужение непрерывного отображения непрерывно.*

Доказательство. Если $f\colon X \to Y$ — непрерывное отображение, а A — подмножество пространства X, то сужение отображения f на A непрерывно как композиция непрерывных отображений — включения $A \to X$ и f:

$$f|_A = f \circ \mathrm{in}_A. \quad \square$$

Определение. Образ топологического пространства при непрерывном в целом отображении часто называется *непрерывным образом* этого пространства.

Непрерывность в точке. Технически непрерывность конкретного отображения проще проверять в отдельных точках.

Определение. Отображение $f\colon X \to Y$ называется *непрерывным в точке* $x_0 \in X$, если для всякой окрестности V образа $f(x_0)$ этой точки существует окрестность U самой точки x_0, образ которой содержится в выбранной окрестности V:

$$f(U) \subset V.$$

В случае, когда X и Y — метрические пространства, в качестве U и V достаточно рассматривать шаровые окрестности; поэтому в этом случае наше определение непрерывности в точке равносильно классическому определению из математического анализа:

Отображение $f\colon X \to Y$ *непрерывно в точке* $x_0 \in X$, если для любого положительного числа ε существует такое положительное число δ, что образ любой точки $x \in X$, удаленной от x_0 менее чем на δ, удален от образа точки x_0 менее чем на ε:

$$\operatorname{dist}(x, x_0) < \delta \Rightarrow \operatorname{dist}(f(x), f(x_0)) < \varepsilon.$$

Вернемся к топологическим пространствам.

Теорема 3. *Отображение $f\colon X \to Y$ непрерывно в целом тогда и только тогда, когда оно непрерывно в каждой точке пространства X.*

Доказательство. Пусть отображение f непрерывно в целом, а $x_0 \in X$ — произвольная точка. Проверим определение непрерывности в точке. Для произвольной окрестности V точки $f(x_0)$ в качестве искомой окрестности U можно взять множество $f^{-1}(V)$: оно содержит точку x_0 и открыто в силу непрерывности в целом отображения f. Наконец, включение $f(U) \subset V$ очевидно.

Докажем обратное утверждение.

Пусть отображение f непрерывно в каждой точке. Нам нужно проверить, что прообраз произвольного открытого множества $V \subset Y$ открыт в X. Для этого заметим, что каждая точка x_0 множества $f^{-1}(V)$ внутренняя, т. е. лежит в нем вместе с некоторой своей окрестностью U. Наличие такой окрестности следует из непрерывности отображения f в точке x_0 и того, что V является окрестностью точки x_0. □

Всюду в дальнейшем мы будем называть непрерывные в целом отображения просто непрерывными.

§ 6. Гомеоморфизмы

Определение. Отображение одного топологического пространства в другое называется *гомеоморфизмом*, если оно непрерывно, обратимо и обратное к нему отображение тоже непрерывно[25].

Таким образом, если X и Y — топологические пространства, то непрерывное отображение $f\colon X \to Y$ является гомеоморфизмом, если f обратимо (т. е. f — биекция) и отображение $f^{-1}\colon Y \to X$ также непрерывно.

Из непрерывности и обратимости отображения еще не следует непрерывность обратимого отображения.

Контрпримеры. 1. Пусть $X = \{a, b\}$ и $Y = \{c, d\}$ — двухточечные топологические пространства. Если пространство X дискретно, а пространство Y антидискретно, то отображение $f : X \to Y$, $a \longmapsto c$, $b \longmapsto d$ очевидно является непрерывным и обратимым. В то же время обратное отображение не будет непрерывным.

2. Рассмотрим хорошо знакомое нам отображение $f\colon [0, 2\pi) \to \mathrm{S}^1 \subset \mathbf{R}^2$ полуоткрытого интервала в единичную окружность, заданное формулой $f(t) = (\cos t, \sin t)$. Оно непрерывно и биективно. Но обратное отображение терпит разрыв в точке $(1, 0) \in \mathrm{S}^1$.

Примеры гомеоморфизмов.

1. Биективное отображение дискретного пространства в дискретное всегда является гомеоморфизмом. Это очевидно.

2. Точно так же биективное отображение антидискретного пространства в антидискретное всегда является гомеоморфизмом.

3. Менее тривиальный пример гомеоморфизма:

$$\mathrm{tg}\colon \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbf{R}, \quad x \longmapsto \mathrm{tg}\, x.$$

Обратное отображение здесь — $\mathrm{arctg}\colon \mathbf{R} \to (-\pi/2, \pi/2)$. Сужение этого гомеоморфизма на интервал $[0, \pi/2)$ дает нам гомеоморфизм $[0, \pi/2) \to [0, +\infty)$, а сужение его на интервал $(0, \pi/2)$ дает нам гомеоморфизм $(0, \pi/2) \to (0, +\infty)$.

[25] Иногда, особенно в старых книгах, используется термин «топологическое отображение».

Свойства. Перечислим простейшие, но очень важные свойства гомеоморфизмов.

Теорема 1. а) *Тождественное отображение любого топологического пространства в себя есть гомеоморфизм.*

б) *Отображение, обратное гомеоморфизму, есть гомеоморфизм.*

в) *Композиция двух гомеоморфизмов есть гомеоморфизм.*

Доказательство. Пункты а) и б) очевидны. Докажем пункт в). Пусть $f\colon X \to Y$ и $g\colon Y \to Z$ — гомеоморфизмы. Тогда их композиция $h = g \circ f\colon X \to Z$ непрерывна, так как отображения f и g непрерывны; отображение h — биекция, так как f и g — биекции; наконец, обратное отображение $h^{-1} = f^{-1} \circ g^{-1}\colon Z \to X$ непрерывно, так как f^{-1} и g^{-1} непрерывны. □

Вот еще некоторые свойства гомеоморфизмов.

При гомеоморфизме образ любого открытого множества открыт, а образ замкнутого множества замкнут.

Действительно, пусть $f\colon X \to Y$ — гомеоморфизм, $g = f^{-1}\colon Y \to X$ — обратное отображение и $U \subset X$ — открытое множество. Тогда множество $f(U) = g^{-1}(U)$ открыто в силу непрерывности отображения g (рис. 4).

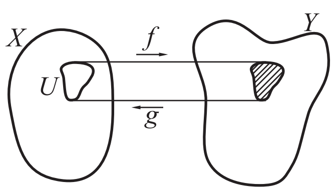


Рис. 4

Замкнутость образа замкнутого множества проверяется аналогично. □

Тем самым *гомеоморфизм* $f\colon X \to Y$ *определяет взаимно однозначное соответствие между топологическими структурами пространств* X *и* Y. Пользуясь этим, нетрудно показать, что если $A \subset X$ — произвольное множество, то

$$f(\operatorname{cl} A) = \operatorname{cl}(f(A)), \quad f(\operatorname{int} A) = \operatorname{int}(f(A)),$$
$$f(\operatorname{fr} A) = \operatorname{fr}(f(A)),$$

а если A есть окрестность точки $x_0 \in X$, то $f(A)$ есть окрестность точки $f(x_0) \in Y$, и т. д.

Таким образом, с топологической точки зрения гомеоморфные пространства устроены совершенно одинаково — гомеоморфизм $X \to Y$ отождествляет все явления в пространствах X и Y, определяемые в терминах топологической структуры.

Здесь мы подошли к одному из важнейших определений курса.

Определение. Говорят, что пространство X *гомеоморфно* пространству Y, если существует гомеоморфизм $f\colon X \to Y$. Гомеоморфность мы будем обозначать значком $\cong$ или, подробнее, $\cong_f$:

$$X \cong Y, \quad X \cong_f Y.$$

Предыдущая теорема 1 позволяет доказать, что:

Теорема 2. *Отношение гомеоморфности является отношением эквивалентности.* Для этого надо проверить три свойства, входящие в определение отношения эквивалентности.

а) Рефлексивность. Всякое топологическое пространство гомеоморфно самому себе:

$$X \cong X.$$

б) Симметричность. Если пространство X гомеоморфно пространству Y, то и пространство Y, в свою очередь, гомеоморфно пространству X:

$$X \cong Y \Rightarrow Y \cong X.$$

в) Транзитивность. Если пространство X гомеоморфно пространству Y, а пространство Y, гомеоморфно пространству Z, то пространство X гомеоморфно пространству Z:

$$X \cong Y,\ Y \cong Z \Rightarrow X \cong Z.$$

Доказательство. Достаточно указать соответствующие гомеоморфизмы. В случае а) это тождественное отображение id_X. В случае б) это $f^{-1}\colon Y \to X$, где $f\colon X \to Y$ — исходный гомеоморфизм. Наконец, в случае в) это $g \circ f\colon X \to Z$, где $f\colon X \to Y$ и $g\colon Y \to Z$ — исходные гомеоморфизмы. При помощи символики это записывается так:

а) $X \cong_{\mathrm{id}} X$;

б) $X \cong_f Y \Rightarrow Y \cong_{f^{-1}} X$;

в) $X \cong_f Y,\ Y \cong_g Z \Rightarrow X \cong_{g\circ f} Z$. $\Box$

Топологический тип. Таким образом, все топологические пространства разбиваются на классы эквивалентности относительно отношения гомеоморфности. Эти классы называются *топологическими типами.* Топологические пространства имеют один и тот же топологический тип в том и только в том случае, если они гомеоморфны.

Важной (особенно в методологическом отношении) является *проблема гомеоморфизма.* Она заключается в том, чтобы определить, гомеоморфны ли два данных топологических пространства. Такая постановка вопроса приобретает смысл, если речь идет о пространствах из

какого-нибудь обозримого класса топологических пространств. Идеальное ее решение заключается в том, чтобы перечислить все топологические типы, встречающиеся в этом классе, и указать способ, позволяющий отнести каждое пространство к своему типу.

Первое составляет содержание *классификации*, второе — *алгоритмической классификации* пространств данного класса.

Топологические свойства. В случае двух пространств положительный ответ на вопрос об их гомеоморфности часто дать проще, чем отрицательный: в простых случаях обычно можно указать конкретный гомеоморфизм одного пространства в другое. Доказать негомеоморфность двух пространств сложнее: для этого надо доказать, что искомого гомеоморфизма *не существует*. На помощь приходят *топологические свойства*. Так называются свойства топологических пространств, которыми гомеоморфные пространства обладают или не обладают одновременно. Примеры топологических свойств: конечность и счетность топологического пространства и топологии, метризуемость, дискретность и антидискретность.

Вообще, топологическим будет любое свойство, которое можно сформулировать исключительно в терминах топологической структуры: открытых и замкнутых множеств.

Топология как раздел геометрии. В XIX в., когда предмет топологии ограничивался множествами в евклидовом пространстве, Ф. Клейн определил топологию как часть геометрии, изучающую свойства фигур, сохраняющиеся при гомеоморфизмах. Это ставило ее в один ряд с другими разделами геометрии, такими как евклидова, гиперболическая, круговая, проективная, аффинная и сферическая геометрии. Их задача, по Клейну, состояла в изучении тех свойств фигур, которые сохраняются, соответственно, при движениях евклидова и неевклидова пространства и круговых, проективных, аффинных и сферических преобразованиях.

Глава II

ТОПОЛОГИЧЕСКИЕ СВОЙСТВА

§ 1. Связность

Определение. Топологическое пространство называется *связным*, если его нельзя разбить на два пустых открытых множества. В противном случае оно называется *несвязным*.

Таким образом, пространство X несвязно, если в нем найдутся два непустых открытых множества U и V, которые не пересекаются, а в объединении дают все пространство X (рис. 5):

$$U, V \in \Omega_X, \quad U \cap V = \varnothing, \quad U \cup V = X.$$

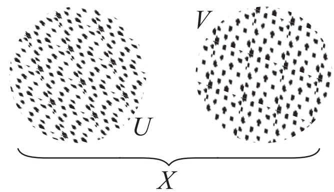


Рис. 5

Можно сказать, что связное пространство «состоит из одного цельного куска»[26].

Поскольку, очевидно, разбиение на два непустых открытых множества одновременно является разбиением на два непустых замкнутых множества, то связность можно переформулировать в терминах замкнутых множеств:

Пространство связно тогда и только тогда, когда его нельзя разбить на два непустых замкнутых множества.

Так как дополнение к множеству непусто тогда и только тогда, когда само множество не совпадает со всем пространством, а дополнение к открытому множеству замкнуто, то мы получаем еще одно полезное определение связности:

Пространство X связно в том и только в том случае, когда в нем имеется только два открыто-замкнутых множества: все пространство X и пустое множество $\varnothing$.

Примеры. 1. Любое антидискретное пространство связно.

2. Дискретное пространство, в котором больше одной точки, несвязно.

3. *Вещественная прямая* **R** *связна* (см. ниже следствие из теоремы 2).

[26]Ср.: связное изложение чего-либо, связный рассказ. Не путать со словом «связанный»!

Определение. Подмножество A топологического пространства X называется *связным*, если оно связно в индуцированной топологии как подпространство, т. е. если топологическое пространство $(A,\ \Omega_A)$ связно.

Другими словами, множество A в топологическом пространстве X связно, если его нельзя покрыть двумя открытыми в X множествами U и V так, чтобы каждое из них пересекалось с множеством A, а пересечение всех трех множеств U, V и A было пусто:

$$U, V \in \Omega_X,\ U \cap A \neq \varnothing,\ V \cap A \neq \varnothing,\ U \cap V \cap A = \varnothing.$$

Примеры. 1. Любое множество в антидискретном пространстве связно.

2. Любое множество в дискретном пространстве, состоящее по крайней мере из двух точек, несвязно.

3. *Интервал* $(0,\ 1) \subset \mathbf{R}$ *связен.* (Доказательство будет дано ниже.)

4. Следующие множества несвязны в $\mathbf{R}$: $[0,\ 1) \cup [2,\ 3]$, $\mathbf{N}$, $\mathbf{Q}$ любое конечное множество. Вообще, если множество $A \subset \mathbf{R}$ содержит точки a и b, но не содержит некоторой промежуточной точки c: $a < c < b$, то оно несвязно. Достаточно положить $U = (-\infty,\ c)$, $V = (c,\ +\infty)$. Таким образом, *если множество в* $\mathbf{R}$ *не является интервалом, оно несвязно.* (Напомним, что *интервалом* называется подмножество прямой $\mathbf{R}$, которое с каждыми своими двумя точками содержит все промежуточные точки. Бывают открытые интервалы, или *промежутки*, конечные вида $(a,\ b)$ и бесконечные вида $(-\infty,\ a)$, $(a,\ +\infty)$, замкнутые интервалы, или *отрезки (сегменты)*, вида $[a,\ b]$ и *полуоткрытые* интервалы вида $[a,\ b)$, $(a,\ b]$, $(-\infty,\ a]$ и $[a,\ +\infty)$. Для полноты к интервалам следует отнести еще пустое множество $\varnothing$, точки — одноточечные множества и всю прямую $(-\infty,\ \infty)$.)

Критерий связности. Достаточный запас связных множеств в топологическом пространстве иногда позволяет судить и о связности всего пространства.

Теорема 1 (критерий связности). *Для того чтобы топологическое пространство было связно, необходимо и достаточно, чтобы каждые две его точки содержались в некотором связном множестве.*

Доказательство. Необходимость этого условия очевидна. Докажем его достаточность. Пусть X — топологическое пространство, каждые две точки которого лежат в некотором связном множестве. Предположим, что пространство X несвязно: $X = U \cup V$, где U и V —

непустые открытые множества, пересечение которых пусто:

$$U \neq \varnothing, \quad V \neq \varnothing, \quad U \cap V = \varnothing.$$

Выберем в этих множествах по точке: пусть $a \in U$, $b \in V$. Пусть $K \subset X$ — связное множество, содержащее выбранные точки a и b. Тогда множества U и V покрывают, очевидно, множество K, их пересечения с K непусты, а пересечения всех трех множеств U, V и K пусто: $U \cap V \cap K \subset U \cap V = \varnothing$. Мы получили противоречие со связностью множества K. Теорема 1 доказана. □

Связность интервалов. Применим полученный критерий связности к вопросу о связных множествах на прямой.

Лемма. *Всякий отрезок* $[a, b]$ *связен.*

Доказательство. Пусть отрезок разбит на два открытых в нем множества: $[a, b] = A \cup B$. Докажем, что если, например, $a \in A$, то множество B пусто. Для этого рассмотрим правый конец α самого большого полуоткрытого интервала $[a, \alpha)$, содержащегося в множестве A. Поскольку множество B открыто, то нетрудно видеть, что $\alpha \notin B$. Следовательно, $\alpha \in A$. Но это означает в силу открытости множества A, что $\alpha = b$ и $A = [a, b]$, а $B = \varnothing$. Лемма доказана.

Теорема 2. *Подмножество A числовой прямой связно в том и только в том случае, когда A — интервал.*

Доказательство. Необходимость этого условия нам уже известна. Чтобы доказать его достаточность, воспользуемся критерием связности и леммой. Действительно, если A — интервал, то каждые две его точки a, b содержатся в связном множестве $[a, b]$.

Следствие. *Вещественная прямая* $\mathbf{R}$ *связна.* □

Связные множества. Рассмотрим теперь некоторые свойства связных множеств.

Теорема 3. *Замыкание связного множества связно.*

Доказательство. Рассмотрим в топологическом пространстве X связное множество A. Предположим, что его замыкание $\operatorname{cl} A$ несвязно. Пусть $\operatorname{cl} A \subset U \cup V$, где U и V — открытые множества, которые имеют непустые пересечения с множеством $\operatorname{cl} A$, причем пересечение всех трех множеств U, V и $\operatorname{cl} A$ пусто:

$$U \cap \operatorname{cl} A \neq \varnothing, \quad V \cap \operatorname{cl} A \neq \varnothing, \quad U \cap V \cap \operatorname{cl} A = \varnothing.$$

Так как множество A связно, то его пересечение с одним из двух множеств U и V пусто. Пусть, например, $U \cap A = \varnothing$. Это значит, что множество A содержится в *замкнутом* множестве $X \setminus U$. Но тогда

и его замыкание $\operatorname{cl} A$ тоже содержится в $X \setminus U$, т. е. $\operatorname{cl} A \cap U = \varnothing$. Полученное противоречие и доказывает теорему 3. $\square$

Теорема 4. *Объединение двух связных множеств, имеющих по крайней мере одну общую точку, связно.*

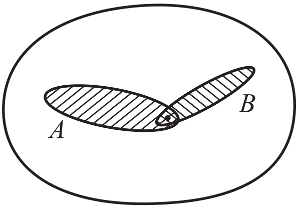


Рис. 6

Доказательство. Пусть A, B — связные подмножества топологического пространства X, имеющие непустое пересечение: $A \cap B \neq \varnothing$ (рис. 6). Предположим, что их объединение $C = A \cup B$ несвязно. Это значит, что $C \subset U \cup V$, где U и V — открытые множества, имеющие общие точки с множеством C, причем пересечение всех трех множеств U, V и C пусто:

$$U \cap C \neq \varnothing, \quad V \cap C \neq \varnothing, \quad U \cap V \cap C = \varnothing.$$

Тогда из связности множеств A и B следует, что каждое из них целиком содержится в одном из множеств U и V и не пересекается с другим (для проверки вернитесь к определению связного множества). Но пусть, например, $A \subset U$. Если теперь $B \subset U$, то $C \cap V = \varnothing$, а если $B \subset V$, то $A \cap B \subset U \cap V \cap C = \varnothing$. В обоих случаях получаем противоречие. Теорема 4 доказана. $\square$

Эту теорему теперь ничего не стоит обобщить на случай произвольного семейства связных множеств, имеющих непустое пересечение.

Теорема 5. *Объединение семейства связных множеств, имеющих общую точку, само связно.*

Доказательство. Пусть $\{A_\alpha\}_{\alpha \in I}$ — произвольное семейство связных множеств в топологическом пространстве X, а точка x_0 — общая для всех множеств A_α: $x_0 \in \bigcap\limits_{\alpha \in I} A_\alpha$. Согласно нашему критерию связности, для доказательства связности множества $\bigcap\limits_{\alpha \in I} A_\alpha$ достаточно указать для двух его точек a и b содержащее обе эти точки связное подмножество множества $\bigcap\limits_{\alpha \in I} A_\alpha$. Если, например, $a \in A_\alpha$ и $b \in A_\beta$ для некоторых индексов α, $\beta \in I$, то таким множеством будет множество $A_\alpha \cup A_\beta$. Его связность следует из предыдущей теоремы, поскольку множества A_α и A_β оба связны и имеют общую точку x_0. $\square$

Компоненты топологического пространства. Особый интерес представляют *максимальные* связные подмножества топологическо-

го пространства. Для них существует специальное название (и даже несколько).

Определение. *Компонентой связности* пространства X называется всякое его связное подмножество, не содержащееся ни в каком строго большем связном подмножестве пространства X.

Теорема 6. *Две компоненты связности либо не пересекаются, либо совпадают.*

Доказательство. Объединение двух пересекающихся компонент по доказанному есть связное множество, которое к тому же содержит обе эти компоненты. По определению, компоненты должны совпадать с ним и, значит, друг с другом. □

Теорема 7. *Каждая точка содержится в некоторой компоненте связности пространства.*

Доказательство. Достаточно заметить, что среди всех связных множеств, содержащих данную точку, существует наибольшее: это объединение всех таких множеств. Оно связно в силу доказанной выше теоремы 5. □

Эти две теоремы в совокупности означают, что всякое топологическое пространство является объединением своих попарно непересекающихся компонент связности. Другими словами, компоненты связности образуют *разбиение* пространства. Компоненты связности пространства называются также его *связными компонентами* или просто *компонентами.*

Теорема 8. *Компонента связности является замкнутым множеством.*

Доказательство. Замыкание компоненты связности есть связное множество, содержащее эту компоненту. По определению, компонента должна совпадать с ним. Следовательно, она замкнута. □

Примеры. 1. У антидискретного пространства, как и у любого связного пространства, всего одна компонента связности — все пространство.

2. В дискретном пространстве каждое одноточечное подмножество образует отдельную компоненту.

3. В множестве рациональных чисел $\mathbf{Q}$, лежащем на вещественной прямой, каждая точка образует отдельную компоненту. Это, кстати, пример того, когда компоненты не открыты.

Связность и непрерывные отображения. Вернемся к связным пространствам. Представляется естественным, что поскольку понятие связности определяется чисто в топологических терминах, то связ-

ность должна быть топологическим свойством. Мы докажем сейчас гораздо более сильное утверждение.

Теорема 9. *Непрерывный образ связного пространства связен. Другими словами, если $f\colon X \to Y$ — непрерывное отображение и пространство X связно, то множество $f(X)$ тоже связно.*

Доказательство. Предположим противное. Пусть множество $f(X)$ несвязно: $f(X) \subset U \cup V$, где U и V — открытые множества в Y, которые имеют непустые пересечения с множеством $f(X)$, причем пересечение всех трех множеств U, V и $f(X)$ пусто:

$$U \cap f(X) \neq \varnothing, \quad V \cap f(X) \neq \varnothing, \quad U \cap V \cap f(X) = \varnothing.$$

Это значит, что $f^{-1}(U) \neq \varnothing$, $f^{-1}(V) \neq \varnothing$ и $f^{-1}(U \cap V) = \varnothing$. В силу непрерывности отображения f множества $f^{-1}(U)$ и $f^{-1}(V)$ открыты, а из того, что $f(X) \subset U \cup V$, следует, что $f^{-1}(U) \cup f^{-1}(V) = X$. Таким образом, мы получили разбиение пространства X на два непустых открытых множества $f^{-1}(U)$ и $f^{-1}(V)$, в противоречие со связностью пространства X. Теорема 9 доказана. □

Следствие. *Топологическое пространство, гомеоморфное связному пространству, само является связным. Таким образом, связность является топологическим свойством.*

Доказательство. Пусть $f\colon X \to Y$ — гомеоморфизм и пространство X связно. Тогда пространство Y связно, как образ связного пространства X при непрерывном отображении f. □

Замечание. Разумеется, поскольку понятие связности определялось в чисто топологических терминах — на языке открытых множеств, — то оно является топологическим. А так как гомеоморфизм устанавливает взаимно однозначное соответствие между топологическими структурами двух пространств, то связность должна сохраняться при гомеоморфизмах, т. е. она является топологическим свойством. Для того читателя, которому все это уже понятно, доказательство предыдущего следствия не нужно. Но, заметим, наше «очевидное» объяснение занимает больше места, чем упомянутое доказательство! В подобной ситуации мы еще не раз окажемся на протяжении этой главы.

§ 2. Линейная связность

Несмотря на то, что определение связности очень просто, его непосредственная проверка в конкретных случаях, например для множеств в евклидовом пространстве, несколько затруднительна. На помощь приходит другое, более сильное, но проще проверяемое топологическое свойство — линейная связность.

Путь. Прежде всего дадим определение пути в топологическом пространстве.

Путем s в топологическом пространстве X называется непрерывное отображение $s\colon\ [0,\,1] \to X$ отрезка $[0,\,1]$ в пространство X. Точка $a = s(0)$ называется *началом* пути s, а точка $b = s(1)$ — его *концом.* Говорят также, что путь s *соединяет* точки a и b (рис. 7)[27].

Определение. Топологическое пространство называется *линейно связным*, если любые две его точки можно соединить путем.

Примеры. 1. Всякое антидискретное пространство линейно связно.

2. Дискретное пространство, в котором больше одной точки, линейно несвязно.

3. *Евклидово пространство* $\mathbf{R}^n$ *линейно связно*: любые две точки $a, b \in \mathbf{R}^n$ можно соединить равномерным прямолинейным путем (фактически

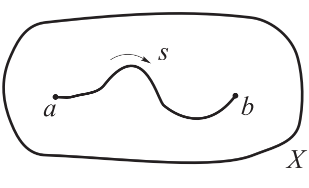


Рис. 7

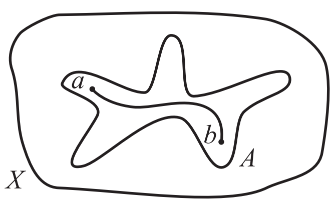


Рис. 8

воспользоваться хорошо известной нам параметризацией отрезка) s : $[0,\,1] \to \mathbf{R}^n$, $s(t) = (1-t)a + tb$.

Определение. Множество A в топологическом пространстве X называется *линейно связным*, если оно линейно связно в индуциро-

---

[27] Отметим некоторую разницу с привычным употреблением слова «путь». Обычно под путем понимают дорогу, а мы имеем в виду скорее *движение* по этой дороге, т. е. прохождение пути.

ванной топологии как подпространство, т. е. если топологическое пространство (A, Ω_A) линейно связно.

Поскольку отображение в подпространство непрерывно тогда и только тогда, когда непрерывна композиция этого отображения с включением подпространства A в пространство X, то мы видим, что *множество A в топологическом пространстве X линейно связно тогда и только тогда, когда любые две его точки a, $b \in A$ можно соединить путем $s\colon [0, 1] \to X$, целиком лежащим в A* (т. е. $s(t) \in A$ для любого $t \in [0, 1]$) (рис. 8).

Линейную связность различных множеств часто нетрудно проверить непосредственно.

Примеры. 1. Любой интервал на прямой линейно связен.

2. Всякое выпуклое множество в евклидовом пространстве линейно связно. В частности, любой шар D^n линейно связен.

Свойства путей. Линейная связность во многом аналогична связности (это видно уже из названия этих свойств). Поэтому неудивительно, что свойства линейно связных множеств очень похожи на свойства связных множеств. Для доказательства этих свойств нам потребуются следующие простейшие свойства путей.

Если точку a можно соединить путем s с точкой b, то и, наоборот, точку b можно соединить с точкой a. Это можно сделать при помощи пути, *обратного* пути s. Он обозначается через s^{-1} и определяется формулой[28]

$$s^{-1}(t) = s(1-t).$$

Можно сказать, что обратный путь s^{-1} состоит в прохождении пути s в обратном направлении (рис. 9).

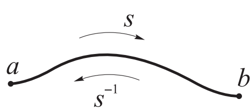


Рис. 9

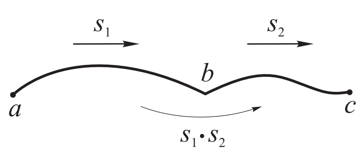


Рис. 10

Если точку a можно соединить путем s_1 с точкой b, а точку b — путем s_2 с точкой c, то точку a можно соединить путем с точкой c. Это

[28]Название «обратный путь» и обозначение s^{-1} несколько двусмысленны, поскольку путь — это отображение, а обратный путь обратным отображением не является. Мы надеемся, что это не вызовет недоразумений.

можно сделать при помощи *произведения путей* s_1 и s_2. Произведение путей обозначается через $s_1 \cdot s_2$ и определяется формулой

$$s_1 \cdot s_2(t) = \begin{cases} s_1(2t), & \text{если} \quad t \leqslant \dfrac{1}{2}, \\ s_2(2t-1), & \text{если} \quad t \geqslant \dfrac{1}{2}. \end{cases}$$

Можно сказать, что путь $s_1 \cdot s_2$ состоит из двух половин — путей s_1 и s_2, каждый из которых проходится с удвоенной скоростью (рис. 10).

Теорема 1. *Объединение семейства линейно связных множеств, имеющих общую точку, само линейно связно.*

Доказательство. Хотя это утверждение и достаточно очевидно, мы его докажем во всех деталях.

Пусть $\{A_\alpha\}_{\alpha \in I}$ — произвольное семейство линейно связных множеств в топологическом пространстве X, а точка x_0 — общая для всех множеств A_α: $x_0 \in \bigcap\limits_{\alpha \in I} A_\alpha$. Как мы знаем, чтобы доказать линейную связность множества $\bigcup\limits_{\alpha \in I} A_\alpha$, нужно для произвольных двух его точек a и b указать соединяющий их путь, который лежал бы в этом множестве. Если, например, $a \in A_\alpha$ и $b \in A_\beta$ для некоторых индексов α, $\beta \in I$, то в силу линейной связности множеств A_α и A_β точку a можно соединить в A_α с точкой x_0 некоторым путем s_1, а точку x_0 можно соединить в A_β с точкой b некоторым путем s_2. Тогда путь $s_1 \cdot s_2$, лежащий в $A_\alpha \cup A_\beta \subset \bigcup\limits_{\alpha \in I} A_\alpha$, соединяет точку a с точкой b и является искомым (рис. 11). $\square$

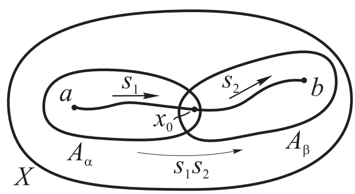


Рис. 11

Компоненты линейной связности топологического пространства. Особый интерес представляют максимальные линейно связные подмножества топологического пространства. Для них имеется специальное название.

Компонентой линейной связности пространства X называется всякое его линейно связное подмножество, не содержащееся ни в каком строго большем линейно связном подмножестве пространства X.

Теорема 2. *Две компоненты линейной связности либо не пересекаются, либо совпадают.*

Доказательство. Объединение двух пересекающихся компонент линейной связности, в силу теоремы 1, есть линейно связное множество, которое к тому же содержит обе исходные компоненты. По определению, они должны совпадать с этим множеством и, значит, друг с другом. □

Теорема 3. *Каждая точка пространства содержится в некоторой его компоненте линейной связности.*

Доказательство. Достаточно заметить, что среди всех линейно связных множеств, которые содержат данную точку, имеется наибольшее: им является объединение всех этих множеств. Оно линейно связно в силу доказанной ранее теоремы 1. □

Теоремы 2, 3 означают в совокупности, что всякое топологическое пространство является объединением своих попарно непересекающихся компонент линейной связности. Другое название для компонент линейной связности — *линейно связные компоненты.*

Линейная связность и отображения. Вернемся к линейно связным пространствам. Нетрудно убедиться, что линейная связность — топологическое свойство. Это естественно, поскольку она определяется чисто в топологических терминах. Мы получим это в качестве следствия из некоторого более сильного утверждения.

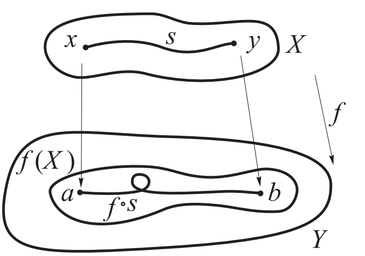


Рис. 12

Теорема 4. *Непрерывный образ линейно связного пространства линейно связен. То есть если f: $X \to Y$ — непрерывное отображение и пространство X линейно связно, то и множество $f(X)$ линейно связно.*

Доказательство. Докажем, что две произвольные точки a, $b \in f(X)$ можно соединить в $f(X)$ путем. Пусть $a = f(x)$, $b = f(y)$, где x, $y \in X$. Тогда если s — путь, соединяющий в X точки x и y, то путь $f \circ s$, очевидно, соединяющий точки a и b, — искомый (рис. 12). □

Пример. Окружность S^1 линейно связна, поскольку она является образом отрезка $[0, 2\pi]$ при стандартной параметризации f: $[0, 2\pi] \to \mathbf{R}^2$, $f(t) = (\cos t, \sin t)$.

Следствие. *Топологическое пространство, гомеоморфное линейно связному пространству, само является линейно связным. Таким образом, линейная связность есть топологическое свойство.*

Доказательство. Если f: $X \to Y$ есть гомеоморфизм линейно

связного пространства X на пространство Y, то пространство Y линейно связно, как образ линейно связного пространства X при непрерывном отображении f. $\square$

Связность н линейная связность. Выясним теперь соотношение между связностью и линейной связностью.

Теорема 5. *Линейно связное топологическое пространство связно.*

Доказательство. Пусть X — линейно связное топологическое пространство. Чтобы доказать, что оно связно, воспользуемся критерием связности. Для этого нужно для произвольных точек a, $b \in X$ найти содержащее их связное множество. Если $s\colon [0, 1] \to X$ — путь, соединяющий точки a и b, то таким множеством будет множество $s([0, 1])$ — «траектория» пути s, — которое связно как образ связного пространства $[0, 1]$ при непрерывном отображении s. $\square$

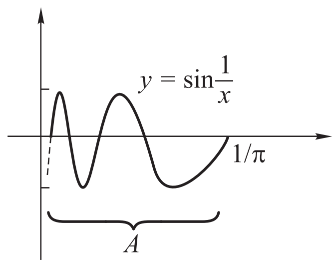


Рис. 13

Обратное утверждение, вообще говоря, неверно: существуют связные, но не линейно связные пространства.

Пример. Рассмотрим непрерывную функцию $f\colon (0, 1/\pi] \to \mathbf{R}$, заданную формулой $f(x) = \sin 1/x$. Пусть $A \subset \mathbf{R}^2$ — ее график. Множество A связно и даже линейно связно. Поэтому его замыкание $\operatorname{cl} A$ (рис. 13) тоже связно. Но нетрудно видеть, что *оно — это замыкание — не линейно связно:* точки отрезка $\operatorname{cl} A \setminus A$ нельзя соединить путем с точками из A (наглядно это довольно очевидно, но мы не будем этого доказывать, а оставим читателю в качестве не очень трудного упражнения).

Заметим, что попутно мы получили пример линейно связного множества, замыкание которого уже не линейно связно. Это еще одно различие между связностью и линейной связностью.

Критерий линейной связности. Бросается в глаза некоторая экзотичность приведенного примера (для пространства $X = \operatorname{cl} A$ есть даже специальное название — «польский отрезок»). Это не случайно. Во многих важных ситуациях связность и линейная связность равносильны. Например, поскольку все интервалы на прямой линейно связны, а все остальные подмножества прямой несвязны, то для подмножеств прямой эти свойства равносильны. Еще более широкий критерий линейной связности связных пространств и множеств дается пунктом

в) следующей теоремы (пункты а) и б) понадобятся нам в § 1 гл. III).

Теорема 6. *Пусть в топологическом пространстве X каждая точка обладает линейно связной окрестностью. Тогда*

а) *компоненты пространства X являются одновременно его линейно связными компонентами;*

б) *компоненты пространства X открыты в нем;*

в) *если пространство X связно, то оно и линейно связно.*

Доказательство основано на следующей лемме.

Лемма. *Всякая линейно связная компонента L пространства X открыта.*

Действительно, все точки множества L внутренние, поскольку с каждой своей точкой множество L содержит ее линейно связную окрестность. □

Доказательство теоремы 6. Начнем с пунктов а) и б). Каждая компонента C пространства X очевидным образом разбивается на линейно связные компоненты. Все они, в силу леммы, открыты. Следовательно, компонента C также открыта, и пункт б) доказан.

Далее, если бы этих линейно связных компонент было больше одной, то мы получили бы противоречие со связностью компоненты C. Следовательно, в разбиении участвует только одна линейно связная компонента L, и она, очевидно, совпадает со всей компонентой C. Этим доказан пункт а).

Пункт в) легко следует из пункта а). Если пространство X связно, то X является (единственной) своей компонентой. Тогда в силу пункта а) пространство X является также своей линейно связной компонентой и, следовательно, линейно связно. Теорема 6 полностью доказана. □

Следствие 1. *Для открытых множеств в евклидовом пространстве связность и линейная связность равносильны.*

Доказательство. Пусть $U \subset \mathbf{R}^n$ — произвольное связное открытое множество. Чтобы доказать его линейную связность, достаточно, в силу теоремы 6, у каждой его точки найти линейно связную окрестность. Такой окрестностью будет, например, любая достаточно малая шаровая окрестность этой точки. □

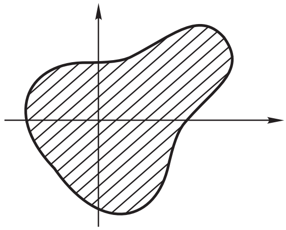


Рис. 14

Напомним, открытое связное множество в евклидовом пространстве называется *областью* (рис. 14).

Следствие 2. *Открытое множество в евклидовом пространстве имеет не более чем счетное число компонент.*

Доказательство остается читателю в качестве упражнения. (Надо воспользоваться, например, тем, что в каждой компоненте должна содержаться точка, у которой все координаты — рациональные числа, а всего таких точек — счетное число.) □

§ 3. Хаусдорфовость

В определении топологического пространства участвуют три аксиомы. Их уже достаточно для довольно содержательной теории. Но запас топологических пространств очень велик. Важнейшие из них — метризуемые — составляют лишь небольшую часть этого запаса. Поэтому часто ограничиваются рассмотрением пространств, удовлетворяющих различным дополнительным требованиям. Среди таких требований одно из важнейших — аксиома Хаусдорфа (иногда ее даже включают в число аксиом топологической структуры).

Определение. Топологическое пространство называется *хаусдорфовым*, если любые две его различные точки обладают непересекающимися окрестностями. Таким образом, если X — хаусдорфово топологическое пространство, то для любых двух точек a, $b \in X$, $a \neq b$, найдутся такие множества U, $V \in \Omega_X$, что $a \in U$, $b \in V$ и $U \cap V = \varnothing$. (рис. 15). (Окрестности U и V *отделяют* точки a и b друг от друга. Поэтому аксиому Хаусдорфа относят к числу *аксиом отделимости*: другие ее названия — вторая аксиома отделимости и аксиома T_2.)

Примеры. 1. Антидискретное пространство, в котором больше одной точки, не хаусдорфово.

2. Дискретное пространство хаусдорфово.

3. Всякое метрическое пространство M хаусдорфово. Если a, $b \in M$ — две его точки, то в качестве отделяющих окрестностей можно взять шаровые окрестности радиусом $\frac{1}{2}\operatorname{dist}(a, b)$.

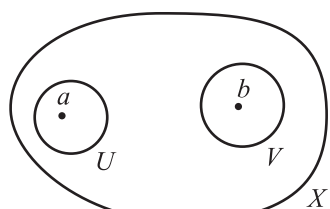


Рис. 15

Первые следствия из аксиомы Хаусдорфа. Из хаусдорфовости следуют некоторые очень естественные свойства топологического пространства.

Теорема 1 (замкнутость точек). *В хаусдорфовом топологическом пространстве X одноточечные подмножества замкнуты.*

Д о к а з а т е л ь с т в о. Пусть $x_0 \in X$. Чтобы доказать, что множество $\{x_0\}$ замкнуто, достаточно проверить, что его дополнение $X \setminus \{x_0\}$ открыто. Действительно, любая точка $y \in X \setminus \{x_0\}$ обладает окрестностью V, не содержащей точки x_0 (в силу хаусдорфовости пространства X), и, следовательно, является внутренней для множества $X \setminus \{x_0\}$. □

Следствие. *В хаусдорфовом пространстве конечные множества замкнуты.*

Д о к а з а т е л ь с т в о. Очевидно. □

Пределы последовательностей. Другое важное следствие хаусдорфовости касается сходящихся последовательностей.

Определение. Пусть в топологическом пространстве X дана последовательность точек $a_1, a_2, \ldots, a_n, \ldots$ Точка a называется (*топологическим*) пределом последовательности $\{a_n\}$, если для любой окрестности U точки a найдется такой номер N, что если $n > N$, то $a_n \in U$. В таком случае говорят, что последовательность $\{a_n\}$ *сходится* к точке a.

Примеры. 1. В антидискретном пространстве любая последовательность сходится к любой точке.

2. Чтобы последовательность $\{a_n\}$ в дискретном пространстве сходилась к точке a, нужно чтобы все ее члены, начиная с некоторого, совпадали с точкой a.

Как обычно, дискретное и антидискретное пространства демонстрируют две крайности. Вообще же, во многих случаях при помощи пределов можно описать большинство топологических явлений, таких как открытость и замкнутость множеств, непрерывность отображений и др. Мы не будем этим заниматься, а ограничимся доказательством следующей теоремы.

Теорема 2 (единственность предела). *В хаусдорфовом пространстве X последовательность $\{a_n\}$ имеет не более одного предела.*

Д о к а з а т е л ь с т в о. Предположим противное. Пусть a и b — пределы последовательности $\{a_n\}$, а U и V — их непересекающиеся окрестности. Тогда, начиная с некоторого достаточно большого номера N, все члены последовательности $\{a_n\}$ будут лежать в U и в V, что, очевидно, невозможно. □

Наследственные топологические свойства. Нетрудно показать, что хаусдорфовость «передается по наследству» от пространства ко всем его подпространствам.

Теорема 3. *Подпространство A хаусдорфова пространства X само является хаусдорфовым.*

Доказательство. Если $a, b \in A$ — две различные точки, а $U, V \subset X$ — их непересекающиеся окрестности в пространстве X, то множества $U \cap A$ и $V \cap A$ будут непересекающимися окрестностями этих же точек в подпространстве A (рис. 16). □

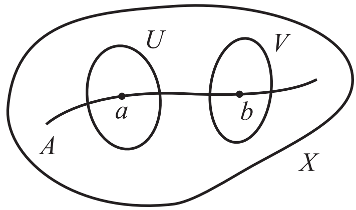


Рис. 16

Топологические свойства, которые передаются таким образом от пространства к его подпространствам, называются *наследственными.*

Пример. Конечность, счетность и метризуемость — наследственные свойства, в отличие, например, от связности и линейной связности.

§ 4. Компактность

Определение. Топологическое пространство называется *компактным*, если всякое покрытие этого пространства открытыми множествами содержит конечное подпокрытие. Покрытия открытыми множествами будем в дальнейшем называть *открытыми.*

Пусть X — топологическое пространство, а $\{U_\alpha\}_{\alpha\in I}$ — некоторое его открытое покрытие: $U_\alpha \in \Omega_X$, $\bigcup\limits_{\alpha\in I} U_\alpha = X$. Тогда наше определение означает, что если пространство X компактно, то среди множеств U_α найдутся несколько множеств: $U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_k}$ (для некоторого $k \in \mathbf{N}$), уже покрывающих его: $X = \bigcup\limits_{i=1}^{k} U_{\alpha_i}$. Множества $U_{\alpha_1}, \ldots, U_{\alpha_k}$ и образуют конечное подпокрытие открытого покрытия $\{U_\alpha\}_{\alpha\in I}$.

Очевидно, для того чтобы пространство не было компактным, у него должно существовать открытое покрытие (заведомо бесконечное), никакая конечная часть которого не является покрытием.

Примеры. 1. Всякое антидискретное пространство компактно.

2. Всякое конечное топологическое пространство компактно.

3. Вообще, всякое пространство, в котором конечное число открытых множеств, компактно.

4. Дискретное пространство с бесконечным числом точек не компактно.

Пример открытого покрытия, не обладающего конечным подпокрытием, дается покрытием одноточечными множествами.

5. Числовая прямая $\mathbf{R}$ не компактна.

Конечным подпокрытием не обладает покрытие всевозможными открытыми промежутками вида (a, b), $a, b \in \mathbf{R}$. Другой пример — покрытие лучами вида $(a, +\infty)$, где $a \in \mathbf{R}$.

С другой стороны, покрытие всеми открытыми лучами $(a, +\infty)$, $(-\infty, a)$ заведомо содержит конечное подпокрытие: $(-\infty, 1) \cup (0, +\infty) = \mathbf{R}$.

Определение. Множество A в топологическом пространстве X называется *компактным*, если оно компактно в индуцированной топологии, как подпространство, т. е. если топологическое пространство (A, Ω_A) компактно.

Теорема 1. *Подмножество A топологического пространства X компактно тогда и только тогда, когда из любого его покрытия множествами, открытыми в X, можно выбрать конечное подпокрытие.*

Доказательство. Начнем с необходимости этого условия. Пусть множество A компактно, а $\{U_\alpha\}_{\alpha\in I}$ — произвольное его покрытие открытыми в пространстве X множествами: $A \subset \bigcup\limits_{\alpha\in I} U_\alpha$, $U_\alpha \in \Omega_X$. Выделим в покрытии $\{U_\alpha\}$ конечное подпокрытие. Для этого рассмотрим пересечения множеств U_α с множеством A: положим $V_\alpha = U_\alpha \cap A$. Множества V_α открыты в подпространстве A и, очевидно, образуют его покрытие: $A = A \cap \bigcup\limits_{\alpha\in I} U_\alpha = \bigcup\limits_{\alpha\in I} V_\alpha$. В силу компактности множества A, из покрытия $\{V_\alpha\}_{\alpha\in I}$ можно выделить некоторое конечное подпокрытие $V_{\alpha_1}, \ldots, V_{\alpha_k}$. Но тогда множества $U_{\alpha_1}, \ldots, U_{\alpha_k}$, очевидно, образуют искомое подпокрытие исходного покрытия множества A: $A = \bigcup\limits_{i=1}^{k} V_{\alpha_i} \subset \bigcup\limits_{i=1}^{k} U_{\alpha_i}$.

Достаточность доказывается аналогично. □

Компактность и замкнутость. Рассмотрим некоторые свойства компактных множеств и пространств.

Теорема 2. *Замкнутое подмножество A компактного пространства X компактно.*

Доказательство. В силу теоремы 1 достаточно из произвольного покрытия $\{U_\alpha\}_{\alpha\in I}$ множества A открытыми в X множествами выбрать конечное подпокрытие. Для этого добавим к этим множествам открытое множество $X \setminus A$ и получим открытое покрытие всего

пространства X. В силу компактности пространства X из этого покрытия можно выделить конечное подпокрытие, причем мы всегда можем считать, что в это подпокрытие входит множество $X \setminus A$. Пусть, например,

$$X = U_{\alpha_1} \cup \ldots U_{\alpha_k} \cup (X \setminus A).$$

Очевидно, что множества $U_{\alpha_1}, \ldots, U_{\alpha_k}$ образуют искомое конечное подпокрытие множества A. $\square$

Однако из компактности множества его замкнутость, вообще говоря, не следует (достаточно вспомнить об антидискретном пространстве, где компактны все множества, а замкнутых — только два). Чтобы это выполнялось, достаточно потребовать хаусдорфовости пространства.

Теорема 3. *Компактное подмножество A хаусдорфова пространства X замкнуто.*

Доказательство. Докажем, что $X \setminus A$ открыто. Для этого достаточно найти для произвольной точки x_0 из дополнения к множеству A окрестность, которая не пересекалась бы с множеством A. В силу хаусдорфовости пространства X у каждой точки $x \in A$ найдется окрестность U_x, не пересекающаяся с некоторой окрестностью V_x точки x_0 (рис. 17).

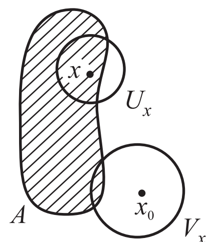


Рис. 17

Всевозможные окрестности U_x, очевидно, образуют открытое покрытие множества A: $A \subset \bigcup\limits_{x\in A} U_x$.

В силу компактности множества A у этого покрытия найдется некоторое конечное подпокрытие: $A \subset U_{x_1} \cup \ldots \cup U_{x_k}$ для каких-то точек $x_1, \ldots, x_k \in A$. Теперь в качестве искомой окрестности точки x_0 можно взять открытое множество $V_{x_1} \cap \ldots \cap V_{x_k}$; оно не пересекается не только с множеством A, но даже с бо́льшим множеством $U_{x_1} \cup \ldots \cup U_{x_k}$. (Действительно, пусть $x \in \bigcap\limits_{i=1}^{k} V_{x_i}$ — произвольная точка. Так как она принадлежит каждому из множеств $V_{x_1}, \ldots, V_{x_k}$, то она не принадлежит ни одному из множеств $U_{x_1}, \ldots, U_{x_k}$. Значит, она не принадлежит и их объединению.) $\square$

Следствие. *Компактное множество в метрическом пространстве замкнуто.*

Доказательство. Достаточно воспользоваться предыдущей теоремой и тем, что всякое метрическое пространство хаусдорфово. $\square$

Компактные множества в метрических пространствах, кроме замкнутости, обладают многими другими специфическими свойствами.

Теорема 4. *Компактное множество A в метрическом пространстве M ограничено.*

Доказательство. Нам нужно указать шар, целиком содержащий множество A. Пусть $a \in M$ — произвольная точка. Всевозможные открытые шары $\mathrm{B}_a(r)$ с центром в этой точке образуют, очевидно, открытое покрытие всего пространства M и, в частности, множества A. Так как множество A компактно, то у этого покрытия найдется конечное подпокрытие. Пусть, например, $A \subset \mathrm{B}_a(r_1) \cup \ldots \cup \mathrm{B}_a(r_k)$ для некоторых $r_1, \ldots, r_k > 0$. Тогда в качестве искомого шара можно взять наибольший из шаров $B_a(r_i)$: ясно, что $A \subset \mathrm{B}_a(R)$, где $R = \max(r_1, \ldots, r_k)$.

Таким образом, *в метрическом пространстве компактные множества замкнуты и ограничены.*

В качестве следствия получаем одно важное свойство компактных множеств на числовой прямой.

Теорема 5. *Компактное подмножество A числовой прямой* $\mathbf{R}$ *содержит свои точные верхнюю и нижнюю грани.*

Доказательство. Так как множество A компактно, то оно ограничено и обладает конечной точной верхней гранью c: $c = \sup A = \sup\limits_{x\in A} x \in \mathbf{R}$. Числа, меньшие числа c, верхней гранью для множества A уже не будут. Это значит, что в сколь угодно малой окрестности точки c имеются точки множества A. Следовательно, точка c является точкой прикосновения множества A. Так как множество A компактно, то оно замкнуто и должно содержать точку c.

Аналогично доказывается, что множество A содержит свою точную нижнюю грань. $\square$

Компактные множества в евклидовом пространстве. По доказанному, все они замкнуты и ограничены. (Этим и объясняется название «компактное» множество.) Оказывается, верно и обратное: из замкнутости и ограниченности множества в евклидовом пространстве следует его компактность. Для доказательства нам потребуется важный частный случай этого утверждения — компактность куба.

Теорема 6. *n-мерный куб в евклидовом пространстве* $\mathbf{R}$ *компактен.*

Доказательство. Рассмотрим для простоты изложения случай плоскости: $n = 2$. Пусть K_0 — единичный квадрат. Докажем от про-

тивного, что множество K_0 компактно. Если это не так, то найдется такое открытое покрытие Γ квадрата K_0, из которого нельзя выделить конечное подпокрытие. Разобьем квадрат K_0 на четыре квадрата со стороной $1/2$. По крайней мере один из них не может быть покрыт конечным числом множеств из покрытия Γ (в противном случае мы нашли бы для каждого из этих квадратов конечное подпокрытие покрытия Γ, и в совокупности эти подпокрытия дали бы конечное покрытие квадрата K_0). Обозначим квадрат, для которого из покрытия Γ нельзя выделить конечное покрытие, через K_1. Разделим и его на четыре квадрата, уже со стороной $1/4$. Тот квадрат, для которого из Γ нельзя выделить конечное подпокрытие, обозначим через K_2. Повторяя эту процедуру, мы получим бесконечную последовательность вложенных квадратов K_0, K_1, K_2, ..., каждый следующий из которых по размерам вдвое меньше предыдущего, и ни один из них не покрывается конечным числом множеств из Γ (рис. 18). Однако общая точка этих квадратов покрывается некоторым множеством из Γ, а с ней должны покрываться этим множеством и все квадраты K_i с достаточно большим i. (Действительно, пусть общая точка лежит в открытом множестве U из Γ вместе со своей ε-окрестностью, где, скажем, $\varepsilon > 1/2^k$. Тогда все квадраты K_i с $i > k$ лежат в этой ε-окрестности и тем более — в множестве U. Противоречие. □

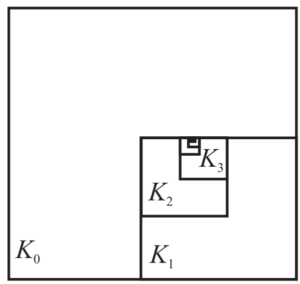


Рис. 18

Следствие. (Критерий компактности в евклидовом пространстве.) *Для того чтобы множество A в евклидовом пространстве $\mathbf{R}^n$ было компактно, необходимо и достаточно, чтобы оно было замкнуто и ограничено.*

Доказательство. Необходимость этих условий нам уже известна. Докажем их достаточность. Так как множество A ограничено, то оно содержится в некотором кубе. А поскольку всякий куб в евклидовом пространстве компактен, то множество A компактно как замкнутое подмножество компактного пространства. □

Примеры. Шары и сферы в евклидовом пространстве компактны.

Компактность и отображения. Вернемся к компактным пространствам. Нетрудно убедиться, что компактность пространства является его топологическим свойством. Имеет место даже более сильное утверждение.

Теорема 7. *Непрерывный образ компактного пространства компактен, т. е. если* $f\colon X \to Y$ — *непрерывное отображение и пространство* X *компактно, то и множество* $f(X)$ *компактно.*

Доказательство. Если открытые множества U_α, $\alpha \in I$, покрывают множество $f(X)$, то их прообразы множества $f^{-1}(U_\alpha)$ покрывают пространство X. В силу непрерывности отображения f это покрытие открыто. Следовательно, из него можно выделить конечное подпокрытие. Пусть, например, $X = f^{-1}(U_{\alpha_1}) \cup \ldots \cup f^{-1}(U_{\alpha_k})$. Но тогда, очевидно, $f(X) \subset U_{\alpha_1} \cup \ldots \cup U_{\alpha_k}$, и множества $U_{\alpha_1}, \ldots, U_{\alpha_k}$ образуют искомое конечное подпокрытие исходного покрытия $\{U_\alpha\}_{\alpha\in I}$. $\square$

Следствие. *Топологическое пространство, гомеоморфное компактному пространству, само является компактным. Таким образом, компактность является топологическим свойством.*

Доказательство. Пусть $f\colon X \to Y$ — гомеоморфизм и пространство X компактно. Тогда пространство Y компактно, как образ компактного пространства X при непрерывном отображении f. $\square$

Попутно мы получаем еще одно следствие из доказанной теоремы. Оно выражает важнейшее свойство компактных пространств и множеств, исключительно полезное при доказательстве различных теорем существования.

Теорема 8. *Непрерывная числовая функция на компактном пространстве ограничена и обладает наибольшим и наименьшим значениями. Другими словами, если* $f\colon X \to \mathbf{R}$ — *непрерывная функция и пространство* X *компактно, то найдутся две такие точки* x_1, $x_2 \in X$, *что* $f(x_1) \leqslant f(x) \leqslant f(x_2)$ *для любой точки* $x \in X$.

Доказательство. Множество $f(X)$ компактно как образ компактного пространства X при непрерывном отображении f. Следовательно, оно содержит свои точные нижнюю и верхнюю грани $\inf(f(X)) = a$ и $\sup(f(X)) = b$.

Пусть $a = f(x_1)$ и $b = f(x_2)$. Тогда, очевидно, точки x_1 и x_2 — искомые. Теорема 8 доказана. $\square$

Критерий гомеоморфизма. Следующая наша цель состоит в доказательстве важного критерия того, когда непрерывное отображение является гомеоморфизмом. Сначала дадим одно определение.

Непрерывное отображение называется *замкнутым*, если при этом отображении образы замкнутых множеств замкнуты. Ясно, что всякий гомеоморфизм есть замкнутое отображение, и наоборот: если непрерывное обратимое отображение замкнуто, то это гомеоморфизм.

Теорема 9. *Непрерывное отображение компактного пространства в хаусдорфово всегда замкнуто. Другими словами, если $f\colon X \to Y$ — непрерывное отображение, пространство X компактно, а пространство Y хаусдорфово, то для любого замкнутого множества $A \subset X$ множество $f(A) \subset Y$ замкнуто.* $\square$

Доказательство. Эта теорема немедленно следует из трех доказанных ранее теорем. Действительно, множество A компактно как замкнутое подмножество компактного пространства X. Множество $f(A)$ компактно как образ компактного множества A при непрерывном отображении f. Наконец, множество $f(A)$ замкнуто как компактное подмножество хаусдорфова пространства Y. $\square$

В качестве следствия мы получаем важный критерий гомеоморфизма.

Теорема 10. *Непрерывная биекция f компактного пространства X на хаусдорфово пространство Y является гомеоморфизмом.*

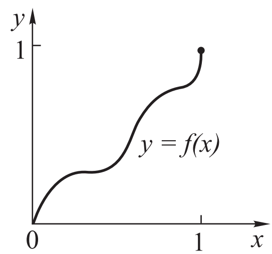


Рис. 19

Доказательство. В силу теоремы 9, отображение f замкнуто, а, как отмечалось пятью абзацами выше, если непрерывная биекция замкнута, то это — гомеоморфизм. $\square$

Пример. *Любая непрерывная биекция отрезка* $[0, 1]$ *в себя есть гомеоморфизм.* Этот факт довольно очевиден наглядно. График отображения в этом случае — кривая без разрывов (рис. 19). И она же даст нам график обратного отображения, если ее отразить относительно прямой $x = y$.

Глава III

МНОГООБРАЗИЯ

§ 1. Топологические многообразия с краем и без края

Понятие многообразия является одним из важнейших в математике. Интересно, что, в сущности, это понятие, а также понятия размерности и края встречаются уже у Евклида. Начнем этот параграф цитатой из Евклида («Начала», Книга I):

Определения.

1. Точка есть то, что не имеет частей.

2. Линия же — длина без ширины.

3. Края же линии — точки.

. . . .

5. Поверхность есть то, что имеет только длину и ширину.

6. Края же поверхности — линии.

Точки, линии и поверхности — все это примеры многообразий. Фактически многообразие — это такое топологическое пространство, в окрестности каждой точки которого можно ввести систему координат. Для более точного определения нам понадобится одно весьма специфическое топологическое свойство — локальная евклидовость. Пусть n — неотрицательное целое число. Через $\mathbf{R}^n_+$ будем обозначать замкнутое полупространство в евклидовом n-мерном пространстве $\mathbf{R}^n$, состоящее из точек, первая координата которых неотрицательна (рис. 20):

$$\mathbf{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \colon x_1 \geqslant 0\}.$$

(Заметим, что $\mathbf{R}^0_+ = \mathbf{R}^0$ состоит из одной точки.)

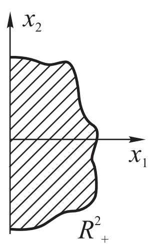


Рис. 20

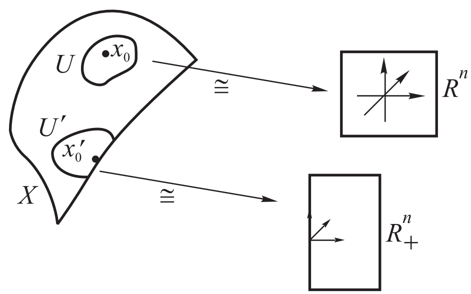


Рис. 21

Определение. Топологическое пространство X называется ***локально евклидовым*** *размерности* n, если всякая его точка x_0 обладает окрестностью U, гомеоморфной n-мерному евклидову пространству $\mathbf{R}^n$ или полупространству $\mathbf{R}^n_+$ (рис. 21). Таким образом, в окрестности U точки x_0 можно ввести систему координат, в которой положение точек, близких к x_0, будет описываться n параметрами. Допуская вольность речи, можно сказать, что вблизи точки x_0 пространство X «имеет n измерений». Поскольку n-мерное евклидово пространство гомеоморфно открытому евклидову шару: $\mathbf{R}^n \cong \mathrm{B}^n$, то в приведенном определении вместо $\mathbf{R}^n$ и $\mathbf{R}^n_+$ можно говорить об открытом шаре и полушаре или просто об открытом множестве в $\mathbf{R}^n_+$.

Примеры. 1. Топологическое пространство является локально евклидовым размерности 0 в том и только в том случае, если оно дискретно.

2. Числовая прямая $\mathbf{R}^1$, окружность S^1, луч $[0, +\infty)$, интервал $[0, 1)$, отрезок $[0, 1]$ являются одномерными локально евклидовыми.

3. Локально евклидовыми размерности n являются сами пространство $\mathbf{R}^n$ и полупространство $\mathbf{R}^n_+$, произвольные их открытые подмножества, n-мерная сфера S^n и n-мерный шар D^n, а также n-мерное проективное пространство.

4. Любое открытое подмножество n-мерного локально евклидова топологического пространства само является n-мерным локально евклидовым.

Локальная евклидовость — очень сильное требование, но само по себе оно не гарантирует выполнения других довольно естественных топологических свойств. Приведем пример, который показывает, что *локально евклидово пространство может не быть хаусдорфовым.*

Пример. Рассмотрим множество X на плоскости, состоящее из всех точек оси абсцисс и еще одной точки — например, точки с координатами $(0, 1)$, рис. 22: $X = \mathbf{R}^1 \cup \{(0, 1)\} \subset \mathbf{R}^2$.[29] Введем в множестве X топологическую структуру: объявим открытыми те множества в X, образы которых при проекции $p : X \to \mathbf{R}^1$ множества X на ось абсцисс открыты в $\mathbf{R}^1$. В этой топологии пространство X локально евклидово размерности 1, поскольку его можно покрыть двумя открытыми множествами, гомеоморфными числовой прямой: $X = \mathbf{R}^1 \cup (X \setminus \{(0, 0)\})$. В то же время пространство X не хаусдорфово: точки $(0, 0)$ и $(0, 1)$ не обладают в нем непересекающимися окрестностями. Действительно, пусть U и V — произвольные окрестности этих точек. Тогда $p(U)$ и $p(V)$ — окрестности точки $0 \in \mathbf{R}^1$. Поскольку множества $p(U)$ и $p(V)$ заведомо имеют общие точки, отличные от 0, а прообраз любой такой точки должен содержаться в $U \cap V$, то и $U \cap V \neq \varnothing$.

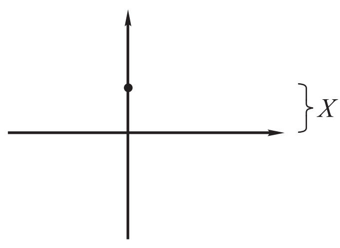


Рис. 22

В следующем основном определении нам будет удобно воспользоваться термином «евклидово множество»: открытое множество в

[29]Здесь $\mathbf{R}^1 = \{(x, y) \in \mathbf{R}^2 \mid y = 0\}$.

n-мерном локально евклидовом топологическом пространстве назовем *евклидовым*, если оно гомеоморфно открытому множеству в полупространстве $\mathbf{R}^n_+$ (т. е. пересечению с $\mathbf{R}^n_+$ открытого множества в $\mathbf{R}^n$).

Основное определение. Топологическое пространство X называется *n-мерным топологическим многообразием*, если оно:

а) локально евклидово размерности n;

б) хаусдорфово;

в) обладает покрытием, состоящим из не более чем счетного числа евклидовых открытых множеств.

Условие в) означает, что пространство X не слишком велико. Часто его заменяют другими условиями: так называемыми сепарабельностью или второй аксиомой счетности. Условие в) автоматически выполнено, например, для компактных локально евклидовых пространств. Без доказательства отметим, что оно гарантирует метризуемость пространства X и даже вложимость его в евклидово пространство достаточно большой размерности.

В дальнейшем мы часто будем называть топологические многообразия просто *многообразиями.*

Примеры. 1. Всякое не более чем счетное дискретное пространство является нульмерным топологическим многообразием. В то же время несчетное дискретное пространство (например: множество вещественных чисел, наделенное дискретной топологией) топологическим многообразием не является. Для него выполнены условия а) и б) определения, но не выполнено условие в): всякое евклидово открытое множество в дискретном пространстве состоит из одной точки, и в силу несчетности пространства никакой счетный набор таких множеств покрытия не образует.

2. Пространство $\mathbf{R}^n$, полупространство $\mathbf{R}^n_+$, их открытые подмножества, сфера S^n, шар D^n — все являются n-мерными топологическими многообразиями. То, что они локально евклидовы, мы уже отмечали. Хаусдорфовость их очевидна. Наконец, что касается их евклидовых покрытий, то $\mathbf{R}^n$, $\mathbf{R}^n_+$ и их открытые подмножества сами образуют одноэлементные такие покрытия, а сфера S^n и шар D^n могут быть покрыты двумя евклидовыми множествами: дополнениями «северного полюса» и «южного полюса» соответственно.

3. Открытое множество в n-мерном топологическом многообразии само является n-мерным топологическим многообразием.

Компоненты топологических многообразий. Поскольку всякое многообразие локально евклидово, то оно очевидным образом ло-

кально линейно связно. Теперь из теоремы 6, § 2 гл. II следует такой результат.

Теорема 1. а) *Компоненты топологического многообразия являются одновременно его линейно связными компонентами.*

б) *Компоненты n-мерного многообразия открыты в нем и, следовательно, являются n-мерными многообразиями.* □

Следующая теорема ограничивает возможное число компонент.

Теорема 2. *Многообразие X состоит из не более чем счетного числа компонент. Компактное многообразие имеет лишь конечное число компонент.*

Доказательство. По определению, X можно покрыть счетным числом евклидовых открытых множеств $X = \bigcup\limits_{i=1}^{\infty} U_i$. Каждое из открытых множеств U_i, будучи евклидовым, имеет, самое большее, счетное число компонент (см. гл. II, § 2, второе следствие из теоремы 6). Поскольку объединение счетного числа не более чем счетных множеств счетно, то мы получаем, что X можно покрыть счетным числом связных множеств: $X = \bigcup\limits_{j=1}^{\infty} A_j$. Далее, так как каждая компонента пространства X содержит по крайней мере одно из множеств A_j, причем разные компоненты, очевидно, содержат разные множества, то компонент не может быть больше, чем множеств A_j; т. е. число компонент не более чем счетно.

Пусть теперь многообразие X компактно. Поскольку его компоненты образуют, очевидно, открытое покрытие пространства X, то X можно, в силу компактности, покрыть конечным числом компонент. С другой стороны, в таком подпокрытии должны участвовать все компоненты. Следовательно, их число конечно. □

Информация без доказательства. Существуют связные (заведомо некомпактные) локально евклидовы хаусдорфовы пространства, не являющиеся многообразиями (они не допускают счетного покрытия открытыми евклидовыми множествами).

Инвариантность размерности многообразия. Остановимся подробнее на понятии *размерности* многообразия. Вполне осмысленным является вопрос: *может ли многообразие размерности n быть одновременно многообразием размерности m при $m \neq n$?* Ответ на него отрицательный. Таким образом, размерность многообразия является его топологическим инвариантом. Для доказательства надо заметить, что в противном случае мы нашли бы у некоторой точки многообразия окрестность, гомеоморфную одновременно открытому мно-

жеству в $\mathbf{R}^m$ и открытому множеству в $\mathbf{R}^n$. Это невозможно в силу следующей теоремы Брауэра, которую также называют «теоремой об инвариантности области». Она утверждает, что свойство множества в евклидовом пространстве быть областью топологически инвариантно, другими словами — это топологическое свойство.

Теорема 3 (Брауэра об инвариантности области). *Если множества A и B в евклидовом пространстве $\mathbf{R}^n$ гомеоморфны и множество A открыто, то и множество B открыто.*

Следствие. *Евклидовы пространства разных размерностей негомеоморфны*: если $m \neq n$, то $\mathbf{R}^m \ncong \mathbf{R}^n$.

Теорема Брауэра бессодержательна при $n = 0$ и несложно доказывается при $n = 1$. Уже при $n = 2$, а тем более при $n \geqslant 3$ ее доказательство представляет значительные трудности, хотя и может быть совершенно элементарным.

Доказательство следствия. Пусть, для определенности, $m < n$. Положим, в условиях теоремы 3, $A = \mathbf{R}^n$ и $B = \mathbf{R}^m$. Тогда теорема 3 утверждает, что $\mathbf{R}^m$ открыто в $\mathbf{R}^n$ (?!). $\square$

Край топологического многообразия. Точка x_0 в n-мерном многообразии называется *внутренней,* если она обладает окрестностью, гомеоморфной евклидову пространству $\mathbf{R}^n$ (не путать с внутренними точками множества в топологическом пространстве!). Точка x_0 называется *краевой*, если у нее существует окрестность U, гомеоморфная полупространству $\mathbf{R}^n_+$, причем связывающий их гомеоморфизм $f\colon U \to \mathbf{R}^n_+$ переводит точку x_0 на границу полупространства $\mathbf{R}^n_+$ (ср. рис. 21). Краевые точки многообразия X образуют его *край*, который обозначается через ∂X. Многообразия, все точки которых являются внутренними, называются *многообразиями без края*, а многообразия, у которых есть краевые точки, называются *многообразиями с краем.* Многообразие без края называется *замкнутым*, если оно компактно, и называется *открытым*, если у него нет компактных компонент.

Примеры. 1. Все точки в $\mathbf{R}^n$ и S^n являются внутренними. В $\mathbf{R}^n_+$ и D^n точки, не лежащие на границе, являются внутренними, а точки, лежащие на границе — краевыми. Таким образом, $\mathbf{R}^n$ и S^n — многообразия без края, причем S^n — замкнутое, а $\mathbf{R}^n$ — открытое, а $\mathbf{R}^n_+$ и D^n — многообразия с краем.

2. Примерами двумерных многообразий без края являются поверхности в $\mathbf{R}^n$: кроме сферы это тор, *крендель* (все это замкнутые многообразия), однополостный и двуполостный гиперболоиды (это открытые многообразия) и т. п. (рис. 23).

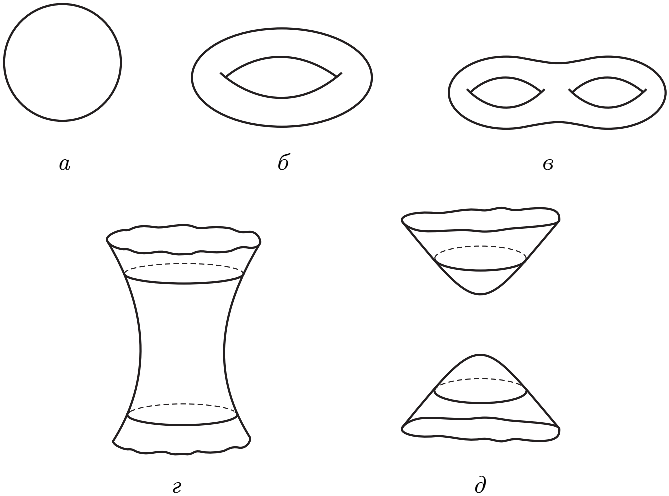


Рис. 23

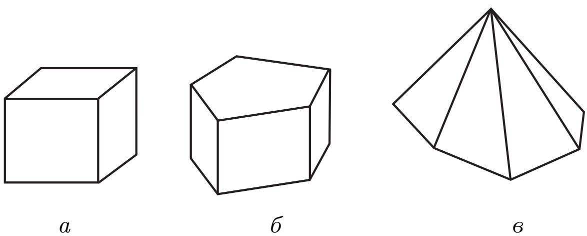


Рис. 24

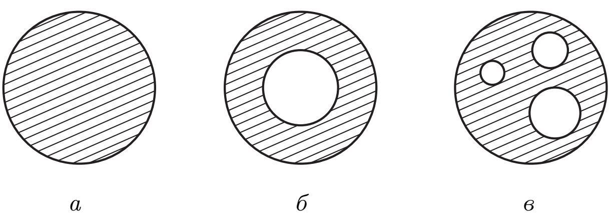


Рис. 25

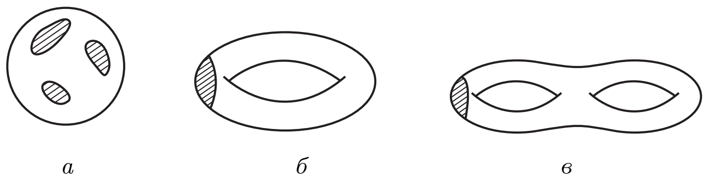


Рис. 26

Особо стоит отметить выпуклые многогранные поверхности, такие как куб, призма, пирамида и т. д. Все они гомеоморфны сфере и тем самым являются замкнутыми двумерными многообразиями (рис. 24).

3. Примерами двумерных многообразий с краем служат: а) замыкания различных плоских областей: кроме круга это кольцо, круг с дырами и т. п. (рис. 25); б) замыкания различных открытых множеств в двумерных многообразиях без края: «сфера с дырами», «тор с дырами», «крендель с дырой» и т. п. (рис. 26).

4. Интересным примером двумерного многообразия с краем в $\mathbf{R}^3$ является так называемый *лист Мёбиуса*. Он выглядит как результат склеивания концов перекрученной полоски бумаги (рис. 27). Лист Мёбиуса — простейшая *односторонняя* поверхность. Что это значит? Обычно у поверхности две стороны: вы можете покрасить одну сторону, скажем, в синий цвет, а другую — в красный, так, что цвета нигде не будут граничить друг с другом. Начав же красить с любого места лист Мёбиуса, вы непременно закрасите его целиком — «со всех сторон»! Две стороны исходной полоски бумаги отождествились при склеивании.

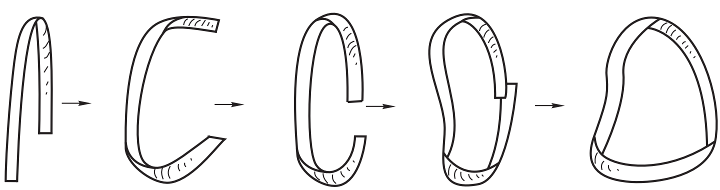


Рис. 27

Инвариантность края многообразия. Наше определение внутренних и краевых точек может вызвать резонный вопрос: *не может ли внутренняя точка многообразия одновременно являться краевой*? (Другими словами, не может ли многообразие без края иметь непустой

край?) Он сводится к вопросу о том, существует ли у точки, лежащей на границе полупространства $\mathbf{R}^n_+$, окрестность в $\mathbf{R}^n_+$, гомеоморфная евклидову пространству $\mathbf{R}^n$. Ответ на этот вопрос отрицательный, как легко следует из теоремы Брауэра об инвариантности области. В частности, край замкнутого полупространства совпадает с граничной гиперплоскостью: $\partial\mathbf{R}^n_+ \cong \mathbf{R}^{n-1}$. Из этого следует, что *край n-мерного многообразия с краем сам является $(n-1)$-мерным многообразием без края.* (Без обращения к теореме Брауэра легко показать, что имеются только две возможности: $\partial\mathbf{R}^n_+ \cong \mathbf{R}^{n-1}$ или $\partial\mathbf{R}^n_+ = \mathbf{R}^n_+$. Во втором случае край любого n-мерного многообразия с ним совпадал бы.)

Примеры. 1. Край отрезка состоит из двух точек — его концов: $\partial[0,\,1] = \{0,\,1\}$.

2. Край круга есть окружность: $\partial\mathrm{D}^2 = \mathrm{S}^1$.

3. Более общим образом, край n-мерного шара есть $(n-1)$-мерная сфера.

4. Край листа Мёбиуса гомеоморфен окружности S^1.

5. Тор является краем ограниченной им части пространства — *полнотория.*

§ 2. Топологические многообразия малых размерностей

В топологии многообразий проблема гомеоморфизма (см. § 6 гл. 1) является центральной. Здесь мы расскажем о ее решении в простейших случаях: опишем топологическую классификацию одномерных и компактных двумерных многообразий (или, иначе говоря, поверхностей). При этом достаточно ограничиться связными многообразиями.

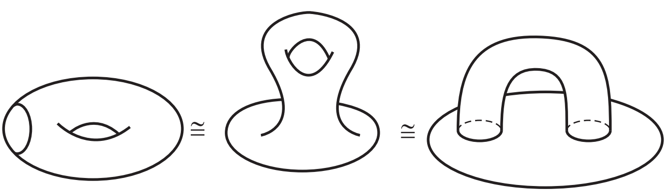


Рис. 28

Теорема 1. *Всякое связное одномерное многообразие без края гомеоморфно числовой прямой $\mathbf{R}^1$ или окружности S^1. Всякое связное одномерное многообразие с непустым краем гомеоморфно отрезку* $\mathrm{I} = [0,\,1]$, *или лучу* $\mathbf{R}^1_+ = [0,\,+\infty)$.

Эту теорему мы доказывать не будем. $\square$

С другой стороны, ясно, что $S^1 \not\cong \mathbf{R}^1$ и $I \not\cong \mathbf{R}^1_+$, — пространства S^1 и I компактны, а $\mathbf{R}^1_+$ и $\mathbf{R}^1$ — нет.

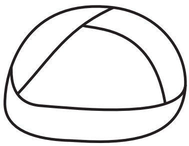


Рис. 29

Ручки, трубки, пленки. Для описания возможных топологических типов замкнутых двумерных многообразий нам понадобятся термины «ручка», «трубка» и «пленка». *Ручкой* будем называть часть двумерного многообразия, гомеоморфную тору с дыркой (рис. 28). *Трубкой* будем называть часть двумерного многообразия, гомеоморфную кольцу (кругу с дыркой). (Название это связано с тем, что трубка гомеоморфна боковой поверхности прямого кругового цилиндра). *Пленкой* будем называть часть двумерного многообразия, гомеоморфную листу Мёбиуса (рис. 29).

Ориентируемость. Назовем двумерное многообразие *неориентируемым*, если оно содержит хотя бы одну пленку, *и ориентируемым* в противном случае.

Сферы с дырами. Напомним, что *сферой с n дырами* называется дополнение в сфере внутренности n попарно непересекающихся кру-

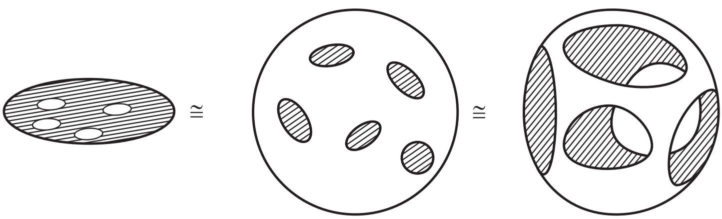


Рис. 30

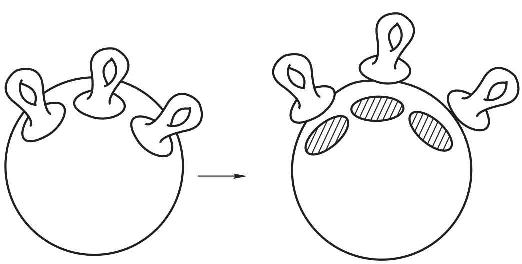


Рис. 31

гов. Сфера с одной дырой гомеоморфна кругу, сфера с двумя дырами гомеоморфна кольцу. Сфера с n дырами гомеоморфна кругу с $n-1$ дырами (рис. 30). *Все сферы с n дырами гомеоморфны.* (Упражнение.)

Сферы с ручками. Замкнутое двумерное многообразие называется *сферой с p ручками*, если в нем можно выделить p попарно непересекающихся ручек, замыкание дополнения к которым гомеоморфно сфере с p дырами (рис. 31). Оказывается, все сферы с ручками ориентируемы. (Этого мы доказывать не будем.)

Все сферы с p ручками гомеоморфны. Вот другое их описание: *ориентируемое двумерное многообразие гомеоморфно сфере с p ручками, если в нем можно выделить p попарно непересекающихся трубок, замыкание дополнения к которым гомеоморфно сфере с $2p$ дырами* (рис. 32).

Примеры. 1. Тор гомеоморфен сфере с одной ручкой (рис. 33).

2. Крендель гомеоморфен сфере с двумя ручками (рис. 34).

3. Еще одна стандартная поверхность, гомеоморфная сфере с p ручками ($p = 4$), изображена на рис. 35 и 36.

Оказывается, что *всякая замкнутая ориентируемая поверхность гомеоморфна сфере с несколькими ручками.* Доказательство этого факта мы отложим до § 4.

Сферы с пленками. Замкнутое двумерное многообразие называется *сферой с q пленками*, если в нем можно выделить q попарно непересекающихся пленок, замыкание дополнения к которым гомеоморфно сфере с q дырами.

Очевидно, что при $q \geqslant 1$ сфера с q пленками не ориентируема. (Так как неориентируемость, по нашему определению, — это наличие на поверхности пленок.)

Пример. Нарисовать сферу с пленками довольно трудно: будучи неориентируемой, она не вкладывается в трехмерное евклидово пространство. Однако *изобразить* ее все-таки можно. Например, вот так выглядит сфера с двумя пленками, иначе называемая бутылкой Клейна (рис. 37, *а*, *б*). На рис. 37, *а* и *б* изображено множество — образ сферы с двумя пленками при некотором ее отображении в $\mathbf{R}^3$ (образы пленок заштрихованы). В этом множестве содержится *двойная окружность*, образованная точками, прообразы которых состоят из двух точек. Бутылка Клейна состоит из двух «половинок», гомеоморфных листу Мёбиуса (см. рис. 38, *а*, *б*).

Труднее изобразить сферу с одной пленкой — хорошо знакомую нам проективную плоскость.

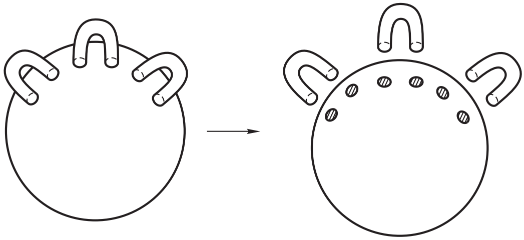


Рис. 32

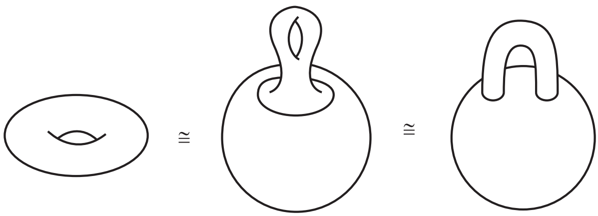


Рис. 33

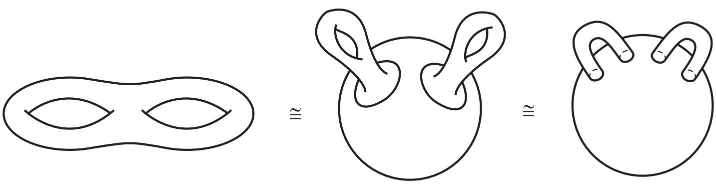


Рис. 34

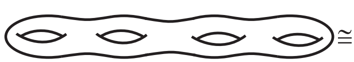


Рис. 35

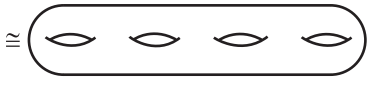
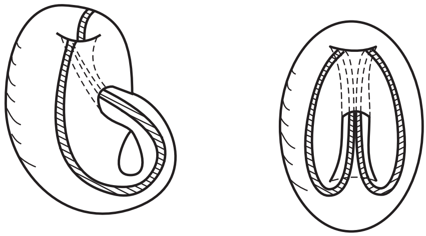
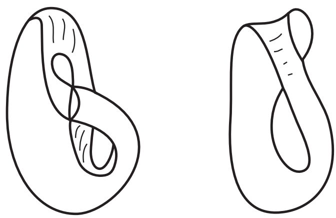


Рис. 36



а *б*

Рис. 37



а *б*

Рис. 38

Без доказательства отметим, что *всякая замкнутая неориентируемая поверхность гомеоморфна сфере с несколькими пленками.*

Модельные поверхности с краем. Назовем поверхность *сферой с p ручками (пленками) и r дырами*, если в ней можно выделить p попарно непересекающихся ручек (пленок), не пересекающихся с краем поверхности, замыкание дополнения к которым гомеоморфно сфере с $p + r$ дырами.

Оказывается, что *всякая ориентируемая компактная поверхность гомеоморфна сфере с несколькими ручками и дырами*, а всякая неориентируемая — сфере с несколькими пленками и дырами.

Для полноты классификации нужно еще доказать, что сфера с пленками и дырами не гомеоморфна сфере с ручками и дырами и что сферы с разным количеством ручек (соответственно пленок) и дыр негомеоморфны. Этого мы здесь делать не будем. Зато в следующем параграфе будет указан способ, позволяющий узнать, какой из модельных поверхностей гомеоморфна данная поверхность.

Многообразия большей размерности. О трехмерных многообразиях известно немало, но до самого последнего времени оставалась не доказанной и не опровергнутой важная гипотеза, выдвинутая еще в начале XX в. французским математиком Анри Пуанкаре. Чтобы сформулировать ее, дадим одно определение. Топологическое пространство X называется ***односвязным***, если оно линейно связно и всякое непрерывное отображение $\mathrm{S}^1 \to X$ окружности в пространство X можно продолжить до непрерывного отображения $\mathrm{D}^2 \to X$ всего круга D^2. Нетрудно видеть, что при $n \geqslant 2$ сфера S^n односвязна.

Гипотеза Пуанкаре. *Всякое замкнутое односвязное трехмерное многообразие гомеоморфно трехмерной сфере.*

Как ни странно, аналоги гипотезы Пуанкаре, касающиеся многообразий размерности 5 и больше, были доказаны первыми. В 1982 г. это было сделано и для многообразий размерности 4. Более того, была получена топологическая классификация вообще *всех* замкнутых односвязных четырехмерных многообразий. Наконец, несколько лет назад исходную гипотезу доказал российский математик Григорий Перельман.

§ 3. Триангуляции, клеточные разбиения. Теорема Эйлера

В результате классификации каждая компактная поверхность должна быть отнесена к какому-то определенному типу. При этом мы получаем представление ее в некотором каноническом виде. Для этого исходная поверхность должна с самого начала быть задана каким-либо способом. Из нашего чисто дескриптивного (т. е. описательного) определения возможный вид такого задания не ясен. Оказывается, что удобно для представления поверхностей воспользоваться *триангуляцией.*

Триангуляции. Пусть F — компактная поверхность. *Топологическим треугольником* T назовем часть поверхности F, для которой установлен гомеоморфизм $\varphi\colon \Delta \to T$ с некоторым плоским треугольником Δ. Образы сторон треугольника Δ назовем *сторонами* треугольника T, а образы вершин — его *вершинами.* Конечный набор треугольников $T_1, \ldots, T_k$ на компактной поверхности F образует ее *триангуляцию*, если выполнены следующие условия:

1) треугольники $T_1, \ldots, T_k$ покрывают поверхность F:

$$\bigcup_{i=1}^{k} T_i = F;$$

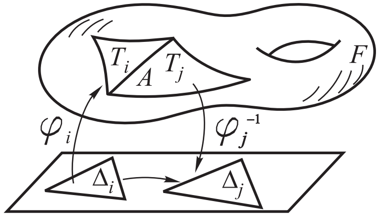


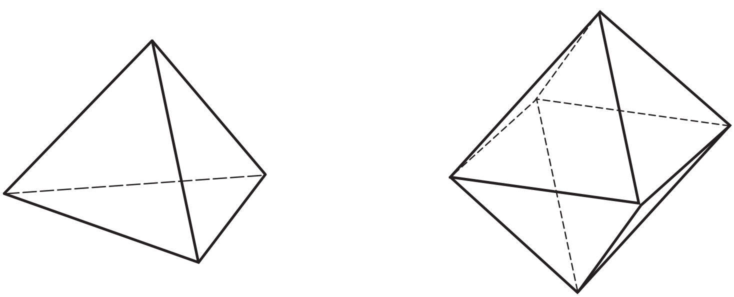
Рис. 39

2) пересечение любых двух треугольников пусто или является их общей вершиной либо стороной;

3) если $A = T_i \cap T_j$ — общая сторона треугольников T_i и T_j, то отображение $\varphi_j^{-1} \circ \varphi_i$ линейно отображает отрезок $\varphi_i^{-1}(A)$ на отрезок $\varphi_j^{-1}(A)$ (рис. 39).

Условие 3 несущественно. Его выполнения всегда можно добиться, изменив гомеоморфизмы φ_i, если условия 1 и 2 уже выполнены.

Примеры. Такие выпуклые многогранники, как тетраэдр и октаэдр, дают триангуляцию двумерной сферы S^2 (см. рис. 40, *а*, *б*). Существует триангуляция тора, в которой участвует всего 14 треугольников. Это — наименьшее возможное число треугольников в триангу-



а *б*

Рис. 40

ляции тора. Пользуясь теоремой Брауэра об инвариантности области (§ 1, теорема 3) при $n = 2$, можно показать, что:

а) три треугольника в триангуляции не могут иметь общую сторону (т. е. одна сторона может принадлежать только одному или двум треугольникам триангуляции);

б) дополнение объединения всех треугольников с данной вершиной до самой этой вершины всегда (линейно) связно.

Эти условия вместе с условиями 1–3 из определения триангуляции можно взять за основу при (чисто комбинаторном) аксиоматическом описании триангулированной поверхности.

Как оказывается (Радо, 1924 г.), *любую компактную поверхность можно триангулировать* (некомпактные поверхности тоже триангулируемы при надлежащем определении триангуляции). Особенно просто это сделать в случае поверхностей геометрического происхождения, которые мы главным образом и имеем в виду. Впервые необходимость триангуляции появилась при проведении измерений на земной поверхности — в геодезии.

Клеточные разбиения поверхностей. Для наших целей удобнее воспользоваться не триангуляцией, а одним ее обобщением — *клеточным разбиением. Открытой клеткой* размерности n на поверхности называется подмножество, гомеоморфное $\mathbf{R}^n$, где $n = 0,\ 1,\ 2$. (Заметим, что открытая клетка размерности n является открытым множеством тогда и только тогда, когда $n = 2$!) Разбиение $F = \bigcup\limits_{i=1}^{k} X_i$ поверхности F на открытые клетки $X_1, \ldots, X_k$ называется *клеточным*, если выполняется следующее условие.

Для каждой одномерной клетки X_i *существует непрерывное отображение отрезка* $\mathrm{I} = [0, 1]$ *в поверхность* F, *сужение которого на интервал* $(0, 1)$ *есть гомеоморфизм интервала* $(0, 1)$ *на* X_i, *а образы концов отрезка* I *являются нульмерными клетками* (*которые, возможно, совпадают*).

Край поверхности при этом тоже получает клеточное разбиение, состоящее из нульмерных и одномерных клеток.

Примеры. 1. Каждая триангуляция поверхности очевидным образом дает ее клеточное разбиение.

2. Выпуклые многогранники, такие как куб, пирамида, призма, додекаэдр и т. п., доставляют клеточные разбиения сферы (рис. 41).

3. Клеточные разбиения бывают намного более «экономными», чем триангуляции. Так, у сферы есть клеточное разбиение, состоящее из одной нульмерной и одной двумерной клетки, а у тора есть разбиение из одной нульмерной, двух одномерных и одной двумерной клетки (рис. 42).

Замкнутые клетки. Замыкание открытой клетки называется *замкнутой клеткой.* В отличие от случая триангуляции замкнутая двумерная клетка может быть не гомеоморфна многоугольнику. Так, в

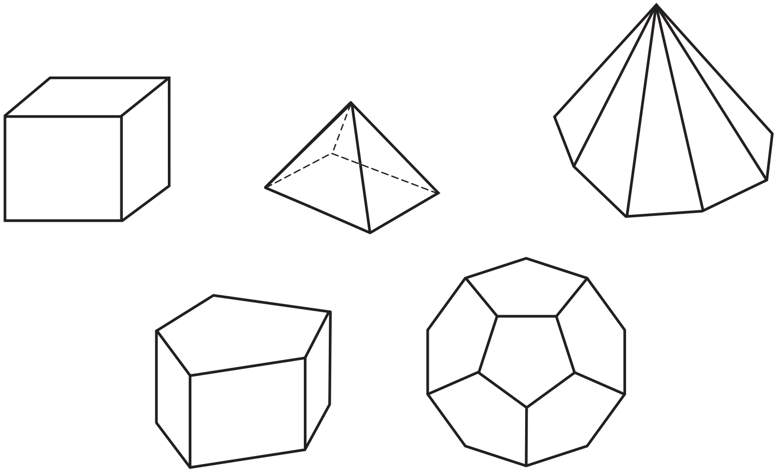


Рис. 41

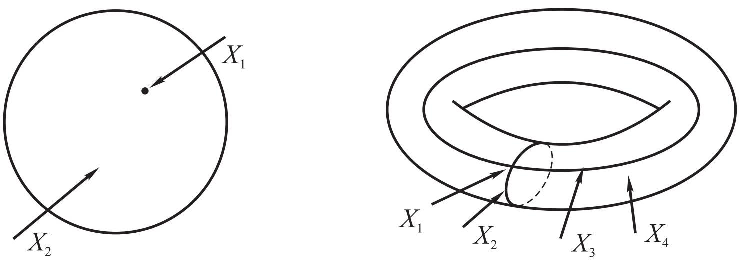


Рис. 42

разбиениях из примера 3 замкнутая двумерная клетка совпадает со всей поверхностью (точно так же и замкнутая одномерная клетка может быть гомеоморфна не отрезку I, а окружности S^1). Тем не менее справедливо некоторое более слабое утверждение. Чтобы его сформулировать, нам понадобится понятие n-угольника при $n = 0$, 1 и 2. Впрочем, определение будет иметь смысл для любого n.

Определение. При $n \geqslant 1$ *модельным n-угольником* назовем круг, на границе которого отмечено n точек (рис. 43, *а–г*). *Нульугольником* назовем сферу с одной отмеченной точкой (рис. 43, *д*).

При $n \geqslant 3$ для модельного n-угольника существует гомеоморфизм его на любой обычный n-угольник, при котором отмеченные точки

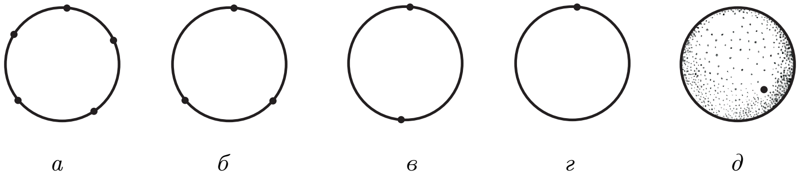


Рис. 43

переходят в вершины, а дуги с отмеченными концами — в стороны. Каждый, модельный n-угольник очевидным образом снабжается клеточным разбиением, в котором участвуют n нульмерных, n одномерных и одна двумерная клетка (исключение представляет нульугольник — у него одна нульмерная, одна двумерная и ни одной одномерной клетки).

Можно показать, что *на всякую замкнутую двумерную клетку можно так отобразить некоторый модельный n-угольник, чтобы каждая его открытая клетка гомеоморфно отображалась на некоторую открытую клетку исходного разбиения* (*той же размерности*). (При этом нульугольник может потребоваться только в том случае, если замкнутая клетка совпадает со всей поверхностью и гомеоморфна сфере S^2.) При желании это свойство можно включить в определение клеточного разбиения поверхности, тем более что в случае триангуляции оно выполнено автоматически.

Пример. Рассмотрим куб $ABCDA'B'C'D'$. Его поверхность, гомеоморфная сфере, обладает клеточным разбиением, в котором нульмерными клетками служат вершины куба, а одномерными — ребра AA', $A'D'$, DD', $D'C'$, $C'C$, $A'B'$, BB'. В этом разбиении всего одна двумерная клетка. Соответствующее отображение многоугольника на замкнутую двумерную клетку (совпадающую со всем кубом) нетрудно построить, взяв за основу хорошо известную развертку куба (рис. 44).

Формула Эйлера. У одной и той же поверхности F может быть

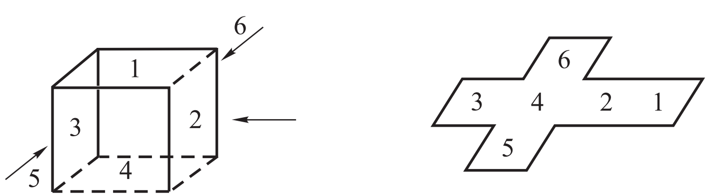


Рис. 44

много различных триангуляций и еще больше клеточных разбиений. Количество клеток в этих клеточных разбиениях тоже может быть самым различным. Но оказывается, что если из общего числа нульмерных и двумерных клеток вычесть количество одномерных клеток, то результат не будет зависеть от выбора разбиения. Полученное число называется *эйлеровой характеристикой* поверхности F и обозначается $\chi(F)$. Если обозначить количество нульмерных, одномерных и двумерных клеток разбиения через α_0, α_1 и α_2 соответственно, то получим

$$\alpha_0 - \alpha_1 + \alpha_2 = \chi(F).$$

Эйлер впервые заметил, что в случае сферы всегда $\alpha_0 - \alpha_1 + \alpha_2 = 2$, т. е. $\chi(\mathrm{S}^2) = 2$. В частности, имеет место

Теорема. *Если дан выпуклый многогранник, у которого* В *вершин,* Р *ребер и* Г *граней, то будет выполняться равенство*

$$\mathrm{В} - \mathrm{Р} + \Gamma = 2.$$

Эта формула была известна еще Ферма и Декарту.

Нетрудно проверить на конкретных примерах клеточных разбиений, что эйлерова характеристика сферы с p ручками и r дырами равна $2-2p-r$, а эйлерова характеристика сферы с q пленками и r дырами равна $2-q-r$. Таким образом, эйлерова характеристика $\chi(F)$, число компонент края r и ориентируемость или неориентируемость связной поверхности F образуют полный набор ее топологических инвариантов.

Доказательство топологической инвариантности эйлеровой характеристики поверхности выходит за рамки теоретико-множественной топологии. Здесь мы ограничимся тем, что докажем формулу Эйлера $\mathrm{В} - \mathrm{Р} + \Gamma = 2$ для выпуклых многогранников.

Доказательство теоремы. Будем последовательно изменять разбиение сферы, соответствующее данному многограннику, стирая некоторые ребра и вершины и следя при этом за изменением чисел α_0, α_1 и α_2. В первый момент $\alpha_0 = \mathrm{В}$, $\alpha_1 = \mathrm{Р}$, $\alpha_2 = \Gamma$. Далее будем поступать так. Каждый раз будем брать какое-нибудь ребро и начинать строить простую ломаную, добавляя к нему по очереди еще не стертые ребра. Если в некоторый момент мы не сможем добавить к ломаной ребро, значит, мы нашли такую вершину, что все выходящие из нее ребра, кроме одного, уже стерты. Тогда сотрем эту вершину и это ребро. При этом соответствующая двумерная клетка изменится (см. рис. 45), но общее количество двумерных клеток останется прежним.

Если обозначить через α'_i количество i-мерных клеток в получившемся разбиении, то мы получим

$$\alpha'_0 = \alpha_0 - 1, \quad \alpha'_1 = \alpha_1 - 1, \quad \alpha'_2 = \alpha_2.$$

Легко видеть, что $\alpha'_0 - \alpha'_1 + \alpha'_2 = \alpha_0 - \alpha_1 + \alpha_2$.

Если же мы сможем построить из нестертых ребер простую замкнутую ломаную, то после стирания любого из входящих в нее ребер разделяемые им двумерные клетки[30] «сольются» в одну (рис. 46), и мы получим: $\alpha'_0 = \alpha_0$, $\alpha'_1 = \alpha_1 - 1$, $\alpha'_2 = \alpha_2 - 1$. По-прежнему $\alpha'_0 - \alpha'_1 + \alpha'_2 = \alpha_0 - \alpha_1 + \alpha_2$.

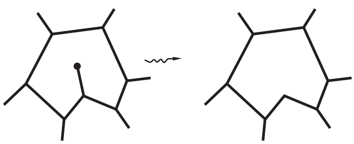


Рис. 45

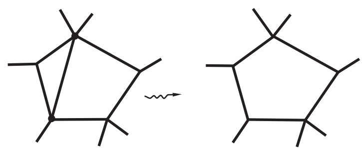


Рис. 46

Повторяя эту операцию, мы в конце концов придем к нульугольнику, у которого $\alpha_0 = 1$, $\alpha_1 = 0$, $\alpha_2 = 1$, и, следовательно, $\alpha_0 - \alpha_1 + \alpha_2 = 2$. Поскольку значение рассматриваемого выражения $\alpha_0 - \alpha_1 + \alpha_2$ при переходе от исходного клеточного разбиения к нульугольнику нигде не менялось, то и для исходного многогранника имеет место равенство $\alpha_0 - \alpha_1 + \alpha_2 = 2$, что и требовалось доказать. $\square$

§ 4. Топологическая классификация ориентируемых замкнутых поверхностей

Наша цель здесь состоит в том, чтобы доказать следующую основную теорему.

Теорема 1. *Всякая замкнутая ориентируемая поверхность гомеоморфна сфере с несколькими ручками.*

Для доказательства воспользуемся очень удобным средством описания поверхностей, которое дают в наше распоряжение клеточные разбиения. Чтобы описать клеточное разбиение поверхности F, изобразим на плоскости (отдельно друг от друга) все многоугольники, соответствующие различным замкнутым клеткам, и пометим у этих

[30] Это клетки, лежащие в разных частях сферы, на которые ее разбивает построенная ломаная.

многоугольников пары сторон, отображающиеся в одну и ту же одномерную клетку. Если одна из сторон в такой паре ориентирована, то сквозной гомеоморфизм определит согласованную ориентацию второй стороны. Такое семейство многоугольников, некоторые стороны которых объединены в пары и при этом согласованно ориентированы, будем называть *выкройкой* поверхности F. Удобно стороны, образующие отмеченную пару, обозначать одной буквой, а ориентацию указывать стрелкой. Выкройку, состоящую из одного многоугольника, назовем *разверткой* поверхности F.

Примеры. На рис. 47 приведена выкройка тетраэдра, соответствующая его триангуляции, на рис. 48 — выкройка сферы с тремя дырами. На рис. 49 изображены развертки известных правильных многогранников, на рис. 50 — развертки трубки, листа Мёбиуса, тора, бутылки Клейна, проективной плоскости (2 штуки). На рис. 51 — развертка сферы с двумя ручками. Аналогично выглядит развертка сферы с произвольным числом ручек. Такая развертка называется *канонической*. На рис. 52 изображена развертка сферы с тремя дырами. Так же выглядит развертка сферы с большим числом дыр.

Замечания. 1. В случае триангуляции удобнее отмечать одной буквой не пары сторон, а вершины, соответствующие одной и той же вершине триангуляции.

2. Можно показать, что всякая выкройка определяет некоторую поверхность с краем — так сказать, результат *склеивания* (или сшивания) кусков выкройки по сторонам, объединенным в пары.

Приступим к доказательству теоремы. Пусть F — произвольная ориентируемая замкнутая поверхность. Рассмотрим какую-нибудь ее выкройку (например, выкройку из треугольников, соответствующую какой-нибудь триангуляции этой поверхности). Занумеруем произвольным образом отмеченные пары сторон многоугольников, образующих выкройку.

Рассмотрим первую отмеченную пару сторон. Если это стороны разных многоугольников выкройки, то выбросим соответствующую им одномерную клетку из разбиения поверхности. При этом две двумерные клетки сливаются в одну. Новому разбиению будет отвечать выкройка, в которой вместо двух прежних многоугольников участвует один, соответствующий новой двумерной клетке (см. рис. 53). Если же эти стороны принадлежат одному и тому же многоугольнику выкройки, то изменим разбиение следующим образом.

Разобьем одномерную клетку, соответствующую нашей паре сторон, на три. При этом придется добавить к разбиению и две одномер-

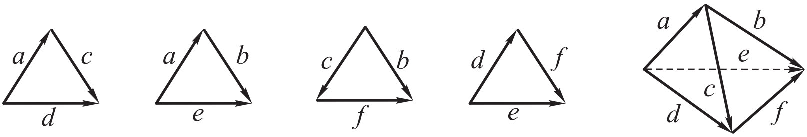


Рис. 47

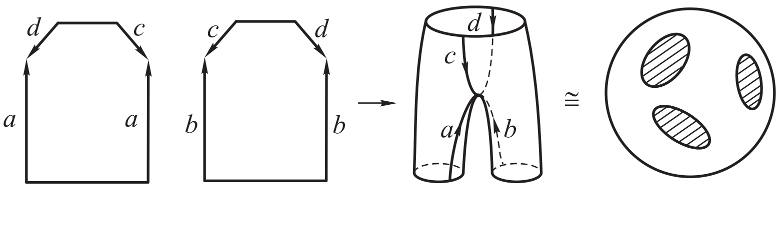


Рис. 48

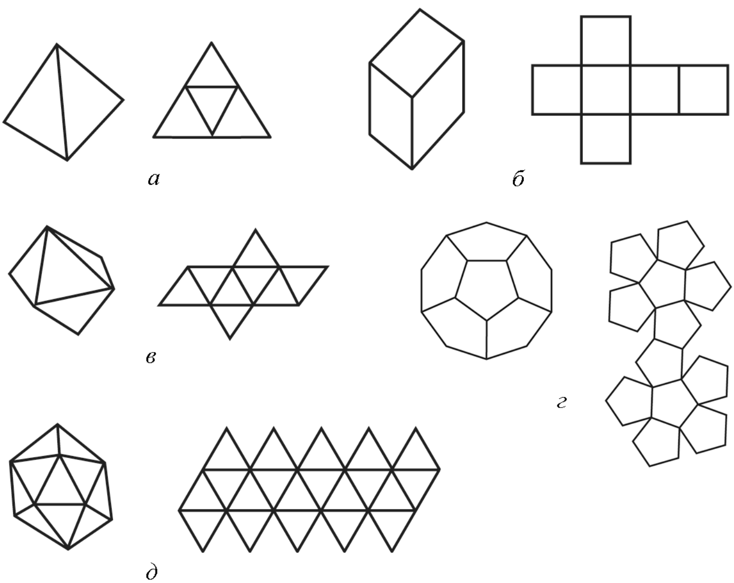


Рис. 49

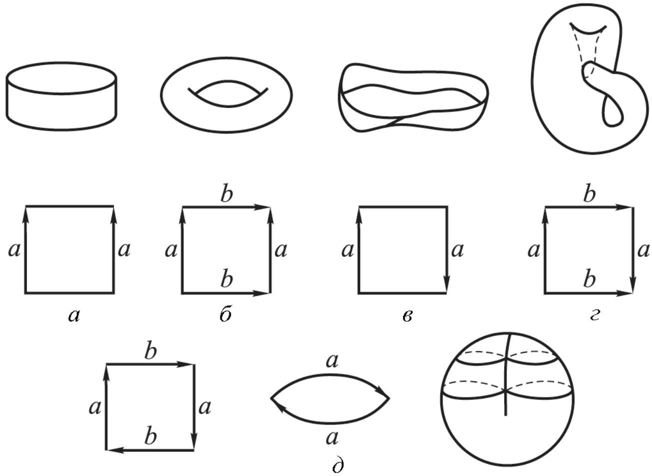


Рис. 50

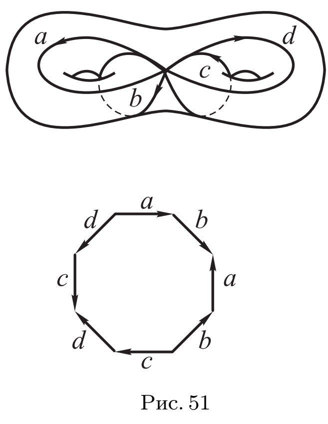


Рис. 51

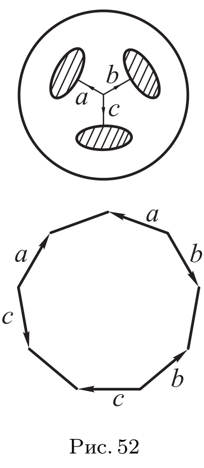


Рис. 52

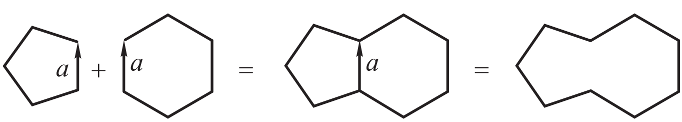


Рис. 53

ные клетки, каждая из которых «начинается и оканчивается» в одной из новых нульмерных клеток. При этом прежняя двумерная клетка разобьется на три новых. Чтобы получить соответствующую выкройку, надо заменить рассматриваемый многоугольник на те три, которые получаются из него при разрезании вдоль двух непересекающихся отрезков, соединяющих отмеченные стороны. При этом образуются пять новых отмеченных пар сторон вместо одной исходной; см. рис. 54.

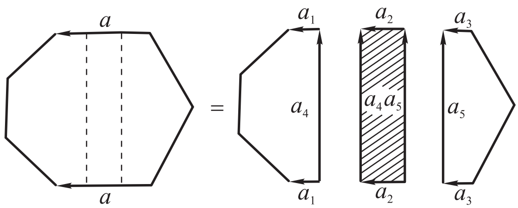
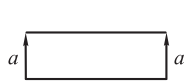


Рис. 54

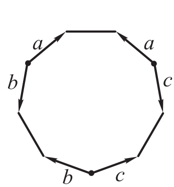
Наш выбор ориентации для сторон, обозначенных буквой a, требует пояснений. Если бы одна из них была ориентирована в другую сторону, то средний четырехугольник, входящий в новую выкройку (на рис. 54 заштрихован) давал бы нам развертку листа Мёбиуса, содержащегося в поверхности F. А это невозможно в силу ее ориентируемости. В частности, отдельно взятый средний четырехугольник образует развертку кольца.

Теперь проделаем то же самое со второй отмеченной парой сторон, потом с третьей и т. д. (При этом мы не будем затрагивать пары сторон, образовавшиеся на предыдущих этапах. Речь идет лишь об исходно занумерованных парах.) По окончании этой процедуры мы получим выкройку исходной поверхности, в которой каждый многоугольник будет либо четырехугольником вида



,

либо иметь вид



.

Соответствующие замкнутые клетки будут гомеоморфны либо кольцу, либо сфере с дырами, причем клетки последнего типа будут пересекаться только с кольцевыми клетками, т. е. с трубками. Таким образом, мы нашли несколько трубок, замыкание дополнения к которым состоит из нескольких сфер с дырами. Для окончания доказательства нам потребуется еще одно описание сфер с ручками.

Предположим, что в ориентируемой поверхности F можно выделить несколько попарно непересекающихся трубок, так чтобы замыкание дополнения к ним состояло из n компонент, каждая из которых была бы гомеоморфна сфере с дырами (рис. 55). Тогда индукцией по n можно показать (как это намечено ниже), что *поверхность F гомеоморфна сфере с ручками.*

Для $n = 1$ это выполняется по определению. Если $n \geqslant 2$, то всегда найдутся две компоненты (на рис. 55 это компоненты A и B) и «соединяющая их» трубка (на рис. 55 — трубка C), объединение которых гомеоморфно сфере с дырами (рис. 56). Ясно, далее, что замыкание дополнения к остальным трубкам (кроме трубки C) состоит уже из $n-1$ компонент, каждая из которых по-прежнему гомеоморфна сфере

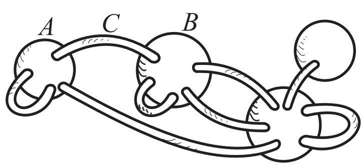


Рис. 55

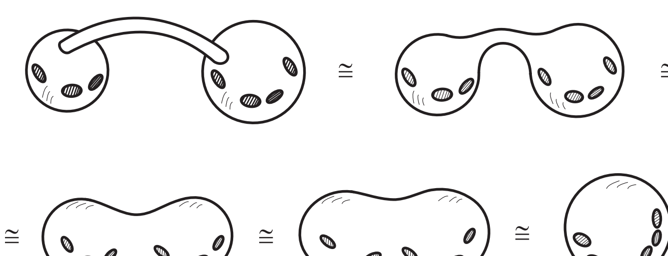


Рис. 56

с дырами. Осталось воспользоваться индуктивным предположением.

Теорема доказана. □

В качестве полезного и интересного упражнения предлагаем читателю проследить все детали доказательства на примере многогранников, как они определены в § 4 гл. II ч. 2. А именно, докажите, что всякая замкнутая многогранная поверхность гомеоморфна сфере с несколькими ручками.

Часть 6
ОСНОВАНИЯ ГЕОМЕТРИИ

В заключительной части, во-первых, даются более глубокие, чем во второй части, основания евклидовой геометрии. А именно, приводится ее аксиоматика, в которой отсутствуют аксиомы о геометрических величинах (длине отрезка, мере угла, площади, объеме). Эта аксиоматика уже не опирается на понятие действительного числа. Решению проблемы измерения геометрических величин посвящены в основном первые две главы этой части. В них же обсуждается важное понятие величины и дается его аксиоматическое определение.

Третья глава посвящена так называемым общим вопросам аксиоматики. Кроме того, в ней даются и сравниваются различные аксиоматики евклидовой геометрии.

Наконец, в последней главе дается обзор различных геометрий, начиная с геометрии Лобачевского и кончая псевдоримановой геометрией — математической основой общей теории относительности.

Глава I
ОСНОВАНИЯ ГЕОМЕТРИИ

Теперь мы войдем в основания геометрии глубже, чем это было сделано в гл. I ч. 2. Там результаты измерения отрезков и углов — численная длина и мера угла — были введены прямо в аксиомах. Здесь же будут даны аксиомы, которые обеспечивают саму возможность измерения отрезков и углов. Кроме того, некоторые аксиомы, сформулированные в гл. I ч. 2, будут заменены другими, содержащими более слабые требования, или разделены на части так, чтобы каждая аксиома выражала только одно условие. (Например аксиома, что две точки соединимы отрезком, и притом только одним, содержит два условия: существование отрезка и его единственность.)

Основные объекты и отношения будут те же, что в гл. I ч. 2.

§ 1. Линейные аксиомы

Основные объекты: 1) *точки*, 2) *отрезки*.

Основные отношения: 1) точка *является концом* отрезка, 2) точка *лежит* на отрезке, 3) два отрезка *равны друг другу* (обозначение: $a = b$).

Если точка M лежит на отрезке a или является его концом, то мы говорим, что она *принадлежит отрезку,* и применяем обычную запись $M \in a$. Соответственно применяем обычные обозначения: $a \subset b$, $c = a \cup b$, где a, b, c — отрезки, и т. п.

I. Аксиомы связи отрезков и концов.

I$_1$. *Для каждого отрезка существуют две точки, являющиеся его концами.*

I$_2$. *Для каждого отрезка существует не более двух точек, являющихся его концами.*

I$_3$. *Для каждых двух точек существует отрезок, концами которого они являются.*

I$_4$. *Существует не более одного отрезка с данными концами.*

Эти аксиомы дают основание к тому, чтобы обозначать отрезок его концами: AB и т. п.

II. Аксиомы о точках на отрезках.

II$_1$. *Для каждого отрезка существует хотя бы одна лежащая на нем точка.*

II$_2$. *Если C на AB, то $AB = AC \cup BC$.*

II$_3$. *Если C на AB, то $AC \cap BC = C$.*

II$_4$. *Если один отрезок содержит конец и еще одну точку другого, то они образуют один отрезок.*

Отрезок не мыслится как множество точек, и потому, строго говоря, нельзя сказать, что объединение $AC \cup BC$ есть отрезок. Равенство $AB = AC \cup BC$ означает, что: (1) каждая точка отрезка AB принадлежит AC или BC, (2) точки отрезков AC, BC принадлежат AB. Такой же смысл имеет утверждение аксиомы II_4: если отрезки a, b удовлетворяют ее условиям, то существует такой отрезок c, что $c = a \cup b$ в указанном выше смысле.

Равенство $AC \cap BC = C$ означает, что у отрезков AC, BC есть единственная общая точка C.

Аксиомы II_2, II_3 выражают, что если C на AB, то отрезок AB составлен из отрезков AC, BC, как это сформулировано в гл. I ч. 2.

Понятно, что аксиомы II_2, II_3 выражают разные условия и поэтому разделены[31].

Условие аксиомы II_4 является частным случаем требования аксиомы I_4 в гл. I ч. 2. Там говорится об отрезках, у которых есть, вообще, две общие точки, здесь же одна из них является концом одного из отрезков. Сказанное в аксиоме I_4 гл. I ч. 2 будет доказано в качестве теоремы.

Укажем основные следствия из высказанных аксиом. Прежде всего напомним три теоремы, доказанные в гл. I ч. 2; их доказательства сохраняются поскольку основаны на тех же аксиомах.

Теорема 1. *Точка, лежащая на отрезке, не является его концом (а потому и конец не лежит на отрезке).* $\square$

Теорема 2. *Если точки C и D принадлежат отрезку AB, то всякая точка, лежащая на CD, лежит также на AB.* $\square$

Теорема 3. *Отрезок определяется принадлежащими ему точками.* $\square$

Теперь две новые теоремы.

Теорема 4. *Если C и D на AB, $C \neq D$, то либо C на AD, либо D на AC (иначе говоря, если C не на AD, то D на AC).*

Доказательство. Пусть C, D на AB, и пусть при этом C не на AD. Докажем, что тогда D на AC. Так как C не на AD, то, по аксиоме деления II_2, C на BD. Допустим, что D на BC, и тем самым, по аксиоме II_2, $BC = BD \cup DC$. Но это противоречит тому, что C на BD. Следовательно, не может быть D на BC, и, значит, по аксиоме II_2, D на AC, что и требовалось доказать. $\square$

Теорема 5. *Если C на AB, а M на AC и N на BC, то C на MN.*

Доказательство. Так как N на BC, то, по теореме 4, C на AN. А так как M на AC, то, по той же теореме, C на MN, что и требовалось доказать. $\square$

Теорема 6. *Если два отрезка a, b имеют две общие точки, то они образуют один отрезок: его концами служат те концы отрезков a, b, которые не лежат ни на a, ни на b.*

Доказательство. Пусть отрезки a, b имеют две общие точки A, B. Если одна из них служит концом одного из этих отрезков, то они образуют общий отрезок, по аксиоме II_4. Если же точка A не является

---

[31] Собственно говоря, равенство $AB = AC \cup BC$, т. е. что отрезок служит объединением двух, выражает два условия: (1) $AB \supset AC \cup BC$ (2) $AB \subset AC \cup BC$. Поэтому аксиому II_2 надо бы еще разделить на две: (а) если C на AB и M на AC, то $M \in AB$, и, меняя обозначения, если M на BC, то $M \in AB$, (б) если C на AB и M на AB, то $M \in AC$ или $M \in BC$.

концом ни одного из отрезков a, b, то она лежит на них обоих. По аксиомам деления $\mathrm{II}_{2,3}$, она делит каждый из них на два отрезка с общим концом A.

Пусть a_1, b_1 — те из этих отрезков, которые содержат точку B, так что у них общий конец A и еще общая точка B. По аксиоме II_4, они образуют один отрезок. Поэтому конец одного из них (отличный от A) содержится в другом. Но это — конец отрезка a или b. Таким образом, для отрезков a и b выполнены условия аксиомы II_4, и, стало быть, они образуют один отрезок.

Концы этого отрезка не лежат ни на a, ни на b, как следует из теоремы 2.

Таким образом, теорема 6 полностью доказана. $\square$

Так же, как в гл. I ч. 2, мы говорим о двух отрезках с общим концом, что один *налегает* на другой, или что они налегают друг на друга, если один из них содержится в другом.

Теорема 7. *Если отрезки AB, AC налегают на один и тот же отрезок AD, то они налегают друг на друга.*

Доказательство. Пусть отрезки AB, AC налегают на AD. Если при этом, например, $AB \supset AD$, то тем самым AB имеет общие точки с AC, кроме A (так как либо $AC \subset AD$, либо $AD \subset AC$). Поэтому, согласно теореме 6, AB налегает на AC.

Остается допустить, что ни AB, ни AC не содержат AD, а стало быть, сами содержатся в AD и не совпадают с AD. Тогда B и C на AD. И, по теореме 4, либо $AB \supset AC$, либо $AC \supset AB$. А это и значит, что отрезки AB, AC налегают друг на друга, что и требовалось доказать. $\square$

III. Аксиомы равенства отрезков.

III_1 (аксиома откладывания). *Для каждых двух отрезков AB, CD существует отрезок AE, равный CD и налегающий на AB.*

III_2 (аксиома меньшего отрезка). *Если отрезок CD содержится в отрезке AB и не совпадает с ним, то он не равен AB.*

III_3 (аксиома сравнения). *Если отрезки равны одному и тому же отрезку, то они равны друг другу.*

III_4 (аксиома сложения). *Если C на AB, C_1 на A_1B_1 и $AC = A_1C_1$, $BC = B_1C_1$, то $AB = A_1B_1$.*

IV. Аксиома непрерывности.

IV_1. *Пусть все точки на отрезке AB разделены на два непустых класса F_1, F_2 так, что если $M \in F_1$ и $N \in F_2$, то $AM \subset AN$. Тогда на AB существует такая точка C, что если $M \in F_1$, то $AM \subset AC$, и если $N \in F_2$, то $AN \supset AC$.*

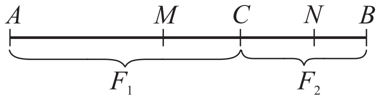


Рис. 1

Поясним наглядно смысл этой аксиомы. Пусть точки внутри отрезка AB разделены на два класса F_1, F_2, как сказано в аксиоме. Представим себе, что точка непрерывно движется по отрезку от конца A к концу B. Она проходит сначала по точкам из F_1, а потом — по точкам из F_2. В каком-то месте должен происходить переход. Это и будет в точке C. Так как $AM \subset AC$ для всех точек M из F_1, то, значит, до C точка движется по точкам M из F_1, а потом — по точкам N из F_2, так как $AN \supset AC$. (К какому классу относится сама точка C — безразлично; ее можно относить к любому из них или к обоим сразу.) Переход происходит в точке C. Если бы ее не было, т. е. аксиома не выполнялась, то при переходе от F_1 к F_2 происходил бы «скачок». Аксиома и выражает, что этого нет, что отрезок «непрерывен» (рис. 1).

Замечание. Среди перечисленных линейных аксиом нет аксиомы, которая обеспечивала бы существование хотя бы одного отрезка и даже хотя бы одной точки. Например, аксиома I_1 говорит, что для каждого отрезка есть две точки, служащие его концами, но это гарантирует существование этих точек, только если существует отрезок. Поэтому все линейные аксиомы имеют условный характер в том смысле, что если есть хотя бы две точки или хотя бы один отрезок, то аксиомы говорят не о пустом множестве... Аксиома, гарантирующая существование точек и отрезков, появится в качестве первой плоскостной аксиомы: существуют три точки, не принадлежащие одному отрезку. С этой аксиомой условный характер линейных аксиом снимается.

§ 2. Алгебра отрезков

В этом и следующем параграфах мы на основе аксиом, высказанных в § 1, обоснуем измерение отрезков. То есть мы докажем в качестве теорем то, что раньше в гл. I ч. 2 было принято в аксиоме измерения. Измерение основано на сравнении и сложении отрезков, а также делении отрезка на равные части (деление масштаба нужно для более точного измерения).

Эти операции и составляют содержание «алгебры отрезков». На ее основе само измерение излагается в следующем параграфе.

Теорема 1. *Отношение равенства отрезков рефлексивно, симметрично и транзитивно, т. е.*: 1) $a = a$, 2) *если* $a = b$, *то* $b = a$, 3) *если* $a = b$ *и* $b = c$, *то* $a = c$.

Доказательство. Равенство отрезков симметрично по понятию, так как говорится: отрезки равны друг другу. Пусть теперь $a = b$ и $b = c$. По симметричности также $c = b$, т. е. получаем, что отрезки a, c равны одному и тому же b, и, стало быть, по аксиоме III_3, $a = c$: равенство транзитивно.

Докажем рефлексивность. Пусть a — данный отрезок. Отложив вдоль него от какого-то его конца отрезок b, ему равный, получим $a = b$, и, по симметричности, $b = a$. По транзитивности: $a = a$. □

Аксиому III_1 откладывания отрезка можно дополнить утверждением о единственности откладываемого отрезка; именно, выполняется

Теорема 2. *При данных отрезках* AB *и* MN *существует не более одного отрезка* AC, *налегающего на* AB *и равного* MN.

Доказательство. Допустим, есть два отрезка AC, AD, налегающих на AB и равных MN. По аксиоме сравнения они равны друг другу, а по теореме 7, § 1 налегают друг на друга, т. е. либо $AC \subset AD$, либо $AD \subset AC$. Но если, скажем, AC содержится в AD и не совпадает с AD, то, по аксиоме III_2 меньшего отрезка, он не равен AD. Получается противоречие. Следовательно, двух отрезков AC, AD быть не может, и теорема доказана. □

Это позволяет формулировать аксиому откладывания III_1 с дополнением о единственности:

Вдоль любого отрезка от любого из его концов можно отложить отрезок, равный данному, и притом только один.

Теорема 3. *Всякий отрезок можно продолжить за любой из концов, т. е. для всякого отрезка* AB *существует отрезок* AC, *составленный из* AB *и* BC. *При этом продолжить можно на отрезок* BC, *равный любому заданному.*

Доказательство. Пусть дан отрезок AB. Возьмем на нем точку D. По аксиоме деления отрезка, $DB \subset AB$. Отложим вдоль DB отрезок DC, равный AB (по аксиоме III_1). Точка C не может принадлежать отрезку DB; иначе получалось бы, что отрезок DC, равный AB, содержится в AB, вопреки аксиоме III_2 о меньшем отрезке.

Отрезки AB, DC имеют общие точки D, B и, следовательно, образуют один отрезок. Это отрезок AC (по аксиоме II_4).

Вдоль отрезка BC можно отложить от точки B отрезок BE, равный любому данному. Вместе с AC он образует (по аксиоме II_4) один

отрезок AE, и отрезок AB оказывается продолженным на отрезок BE, равный данному. □

Сложение отрезков. Отрезок a называется *суммой* отрезков b, c, если он составлен из этих отрезков или из отрезков, им равных; в записи: $a = b + c$ или $a = c + b$, потому что когда отрезок a составлен из отрезков b и c, то порядок их не указывается.

Определенная таким образом операция сложения обладает следующими свойствами.

1. *Суммы равных отрезков равны, т. е. если* $a = b + c$, $a_1 = b_1 + c_1$ *и* $b = b_1$, $c = c_1$, *то также* $a = a_1$. Это утверждается аксиомой сложения III_4.

Поэтому можно считать, что запись $b + c$ обозначает любой такой отрезок a, что $a = b + c$; все такие отрезки равны. □

2. *Любые два отрезка можно сложить, т. е. какие бы не были отрезки* b, c, *существует такой отрезок* a, *что* $a = b + c$.

Действительно, любой отрезок b можно продолжить за любой из концов на отрезок, равный данному отрезку c (теорема 3). □

3. *Если* $a = b + c$ *и* $a_1 = a$, *то* $a_1 = b + c$, *т. е. если отрезок* a *составлен из отрезков* b, c *и* $a_1 = a$, *то отрезок* a_1 *составлен из отрезков, равных* b, c.

Действительно, пусть отрезок $a = AB$ составлен из $b = AC$ и $c = BC$, и пусть отрезок $a_1 = A_1B_1$ равен AB. Отложим вдоль него отрезок A_1C_1, равный AC, а затем приложим к A_1C_1 отрезок C_1B_2, равный CB (рис. 2). Получим отрезок A_1B_2, равный AB (по аксиоме сложения). Он отложен вдоль A_1B_1, а стало быть, совпадает с ним (по теореме 2 о единственности откладывания отрезка). Таким образом, выходит, что $A_1B_1 = A_1C_1 + C_1B_1$, т. е. $a_1 = b + c$, поскольку отрезки A_1C_1 и C_1B_1 равны b и c. □

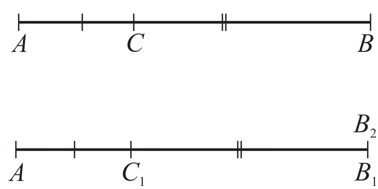


Рис. 2

Доказанное свойство можно выразить так: если точка C на AB и отрезок A_1B_1 равен AB, то на нем есть такая точка C_1, что $A_1C_1 = AC$, $B_1C_1 = BC$.

4. *Сложение коммутативно и ассоциативно* (*переместительно и сочетательно*), т. е.

$$a + b = b + a,$$

$$(a + b) + c = a + (b + c).$$

Первое заключено в определении. Второе очевидно, если отрезки a, b, c обозначать AB, BC, CD. Тогда

$$a + b = AC, \quad (a + b) + c = AC + CD = AD,$$
$$a + (b + c) = AB + (BC + CD) = AB + BD = AD. \quad \square$$

Определение. Если для отрезков a, b существует такой отрезок c, что $a = b + c$, то мы говорим, что a *больше* b или, что то же, b *меньше* a; в записи $a > b$ или $b < a$.

Определение. Мы говорим, что отрезок c есть *разность* отрезков a и b, и пишем $c = a - b$, если $a = c + b$. Если отрезок b отложить вдоль a, то остаток и представляет разность $a - b$ (рис. 3).

Лемма. Если $a > b$ *и* $a_1 = a$, $b_1 = b$, *то также* $a_1 > b_1$ *и* $a_1 - b_1 = a - b$ (*т. е. если от равных отнять равные, то получаются равные*).

Рис. 3

Рис. 4

Доказательство. Пусть $a > b$, так что есть такой отрезок c, что $a = b + c$. Пусть $a_1 = a$, $b_1 = b$. Тогда, по аксиоме сложения, $a = b + c = b_1 + c$. Так как $a_1 = a$, то $a_1 = b_1 + c$, т. е. $a_1 > b_1$ и $a_1 - b_1 = c$. Поэтому $a_1 - b_1 = a - b$, что и требовалось доказать. $\square$

Кратные отрезки и деление отрезка. Если n — натуральное число, то определяем

$$na = \underbrace{a + \ldots + a}_{n \text{ раз}};$$

а также определяем $\frac{1}{n}a$ как такой отрезок b, что $nb = a$, т. е. $n\left(\frac{1}{n}a\right) = a$. Далее, определяем $\frac{m}{n}a = m\left(\frac{1}{n}a\right)$. При этом $m\left(\frac{1}{n}a\right) = \frac{1}{n}ma$ (доказательство этого равенства оставляем читателю).

Теорема 4. *Для всякого отрезка a и всякого натурального n существует отрезок $\frac{1}{n}a$.*

Докажем это для $n = 2$, т. е. что у всякого отрезка AB есть середина — такая точка C, что $AC = CB$.

Для этого заметим следующее: на всяком отрезке AB есть такие точки M, N, что $2AM < AB$ и $2AN > AB$ (рис. 4).

Действительно, возьмем на отрезке AB какую-нибудь точку P. Если $2AP \leqslant AB$ (рис. 5, *а*), то примем за M любую точку на AP. Тогда заведомо будет $2AM < AB$. Допустим, $2AP > AB$, т. е. $PB < AP$, так что $2PB < AB$. Тогда если отложить вдоль AB отрезок AM, равный PB, то и получим $2AM < AB$ (рис. 5, *б*).

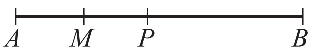


Рис. 5

Если $2AM < AB$, то $2BM > AB$, и поэтому, отложив от A отрезок AN, равный BM, получим, что $2AN > AB$.

Теперь обратимся к нахождению середины какого-либо данного отрезка AB.

Разделим все точки на AB на два класса F_1, F_2: в F_1 отнесем те точки M, для которых $2AM < AB$; в F_2 отнесем все остальные точки, т. е. те N, для которых $2AN \geqslant AB$. Очевидно, при любых таких M, N будет $AM \subset AN$. Из доказанного только что замечания следует, что в обоих классах есть точки.

Таким образом, наши классы F_1, F_2 — такие, о которых говорится в аксиоме непрерывности. По этой аксиоме существует такая точка C, что для всех $M \in F_1$ и всех $N \in F_2$

$$AM \subset AC \subset AN. \tag{1}$$

Отрезок AC и будет половиной AB, т. е. $2AC = AB$

Допустим, $2AC \neq AB$, и пусть $2AC < AB$. По доказанному замечанию можно взять такой отрезок c, что $2c < AB - 2AC$.

Отложим от точки C вдоль CB отрезок CM_1, равный c (рис. 6). Получаем отрезок $AM_1 = AC + CM_1 > AC$, т. е. $AM_1 \supset AC$, и $M_1 \in F_2$.

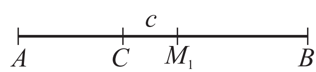


Рис. 6

Вместе с тем (так как $CM_1 = c$ и $2c < AB - 2AC$), $2AM_1 = 2(AC + CM_1) = 2AC + 2c < 2AC + AB - 2AC = AB$, т. е. $2AM_1 < AB$, и, по определению класса F_1, $M_1 \in F_1$. (?!)

Это противоречие доказывает, что не может быть $2AC < AB$.

Допустим, $2AC > AB$. Это равносильно тому, что $2BC < AB$. Поэтому так же, как только что сделано, придем к противоречию. (Именно, укажем такую точку N, что $BN \supset BC$ и вместе с тем $2BN < BA$. Первое равносильно $AN \subset AC$, второе равносильно $2AN > AB$, т. е. $N \in F_2$, так что из (1) $AN \supset AC$. Противоречие.)

Таким образом, существование середины у всякого отрезка и тем самым существование отрезка $\frac{1}{2}a$ доказано.

Существование отрезка $\frac{1}{n}a$ устанавливается аналогично. Но нам оно пока не понадобится, а потому оставляем его здесь без доказательства. Потом существование отрезка $\frac{1}{n}a$ автоматически последует из существования отрезка любой данной длины. □

§ 3. Измерение длины

Сколь угодно большие и сколь угодно малые отрезки. Практический прием измерения длины состоит в откладывании масштаба и его долей. При этом возможность измерить, в принципе, любой отрезок любым масштабом с любой точностью обусловлена двумя фундаментальными обстоятельствами.

(1) Любой отрезок, каким бы большим он ни был, можно перекрыть данным масштабом, откладывая его достаточное число раз.

(2) Любой данный масштаб (отрезок) можно разделить на сколь угодно малые доли, так что ими можно измерять сколь угодно малые отрезки и тем самым измерять любые отрезки с любой точностью.

В геометрии это выражается двумя теоремами.

Теорема 1. *При любых двух отрезках a, b найдется такое натуральное n, что $na \geqslant b$.*

Теорема 2. *При любых двух отрезках a, b найдется такое натуральное n, что $\frac{1}{n}a \leqslant b$* (как доказано, отрезок $\frac{1}{n}a$ существует по крайней мере для $n = 2^k$).

Доказательство теоремы 1. Пусть a, b — данные отрезки. Допустим, вопреки доказываемому, что при всяком n будет $na < b$. Это значит, что если откладывать на отрезке b, начиная от одного его конца A, отрезки, равные a, то будем получать отрезки AA_1, A_1A_2, …, содержащиеся в b. Все их точки, кроме A, лежат на b. Покажем, что существует такой отрезок AC, который содержит все эти отрезки A_iA_{i+1}, причем они исчерпывают все его точки, кроме C.

Может оказаться, что таким будет сам отрезок b. Если же это не так, то на b есть точки, не попадающие в отрезки A_iA_{i+1}. Соответственно, разделим точки на отрезке b на два класса: в класс F_1 отнесем

точки всех отрезков A_iA_{i+1}, в класс F_2 — все остальные точки. Очевидно, эти классы удовлетворяют условиям аксиомы непрерывности. Поэтому, согласно этой аксиоме, есть такой отрезок AC, что он содержит все точки класса F_1 и на нем нет других точек. Так что любой отложенный вдоль CA отрезок CD содержит точки A_i.

Возьмем такой отрезок CD, равный a. Тогда окажется, что отрезок CD, равный a, содержит внутри какой-то отрезок A_iA_{i+1}, тоже равный a. Но это противоречит аксиоме меньшего отрезка.

Следовательно, не может быть, чтобы отрезками na нельзя было покрыть отрезок b. Теорема 1 доказана. $\square$

Доказательство теоремы 2. Пусть a, b — данные отрезки. По теореме 1 существует такое n, что $nb \geqslant a$. А это равносильно тому, что $b \geqslant \frac{1}{n}a$. Теорема 2 доказана. $\square$

Существование длины. Измерение длины путем откладывания масштаба и его долей, представленное в идеализированном виде, приводит к следующему результату.

Теорема 3 (о длине отрезка). *При произвольно выбранном отрезке e каждому отрезку a однозначно сопоставляется число $l(a)$ так, что выполняются условия:*

(а) $l(a) > 0$,

(б) *если* $a = b$, *то* $l(a) = l(b)$,

(в) $l(a + b) = l(a) + l(b)$,

(г) $l(e) = 1$.

Число $l(a)$ называется *численной длиной* отрезка a *в масштабе* e (как говорят: «длина 5 см» и т. п.). «Число» всегда означает действительное число.

Теорема состоит из двух утверждений: существование длины и ее единственность при данном масштабе. Докажем первое.

Теорема 3а. *При данном отрезке e каждому отрезку a можно сопоставить положительное число $l(a)$ так, что будут выполнены условия* (б) — (г).

Доказательство. Пусть выбран отрезок e. Обозначим его доли: $e_n = \frac{1}{2^n}e$.

Определим численную длину отрезков e_n: $l(e) = 1$, $l(e_n) = 1/2^n$; для отрезков ke_n, составленных из e_n, полагаем

$$l(ke_n) = \frac{k}{2^n}. \tag{1}$$

Возьмем какой-либо отрезок a и, по теореме 2, найдем такое n_0, что $e_n \leqslant a$ при $n \geqslant n_0$, так что отрезок e_n укладывается на a. Вместе

с тем, по теореме 1, при любом $n \geqslant n_0$ найдется такое m_n, что

$$a_n = m_n e_n \leqslant a < (m_n + 1)e_n. \tag{2}$$

Если перейти от n к $n+1$, то к отрезку a_n могут прибавиться некоторые отрезки e_{n+1}, содержащиеся в a. Поэтому $l(a_{n+1}) \geqslant l(a_n)$.

Таким образом, последовательность чисел $l(a_n)$ — неубывающая и ограниченная. Следовательно, она имеет предел; мы принимаем его за длину отрезка a:

$$l(a) = \lim_{n\to\infty} l(a_n).$$

Если отрезок a' равен a, то на нем укладывается столько же отрезков, равных e_n, и не более (как это непосредственно следует из выводов «алгебры отрезков»). Стало быть, для a' отрезки a'_n равны a_n. Тем самым

$$l(a') = \lim l(a'_n) = \lim l(a_n) = l(a).$$

Итак, если $a' = a$, то $l(a') = l(a)$.

Пусть теперь $a = b + c$. Отложив на b и c отрезки, равные e_n, получим отрезки b_n, c_n, причем

$$b = b_n + \beta_n, \quad c = c_n + \gamma_n,$$

где β_n, γ_n — «остатки» после откладывания отрезков, равных e_n, так что β_n и γ_n меньше e_n. Не исключено, что отрезков β_n, γ_n нет; можно тогда считать их «нулевыми» отрезками. По правилам алгебры отрезков,

$$b + c = b_n + c_n + (\beta_n + \gamma_n);$$

и так как $\beta_n, \gamma_n < e_n$, то $\beta_n + \gamma_n < 2e_n$. Поэтому если откладывать отрезки e_n на отрезке $a = b + c$, то их уложится максимум на один больше, чем на b и c в сумме, т. е.

$$b_n + c_n \leqslant a_n \leqslant b_n + c_n + e_n,$$

и, переходя к длинам, получаем

$$l(b_n) + l(c_n) \leqslant l(a_n) \leqslant l(b_n) + l(c_n) + \frac{1}{2^n}.$$

Поэтому в пределе при $n \to \infty$ получим

$$l(a) = l(b) + l(c).$$

Таким образом, условие (в) для длины выполнено, и теорема 3а доказана. □

Теорема 4. *Всякие две функции отрезка* l, l', *удовлетворяющие условиям* (а) — (в), *отличаются только постоянным множителем, т. е.* $l'(a) = k \cdot l(a)$, $k = \mathrm{const}$.

Доказательство. Пусть l, l' — функции с условиями (а) — (в). Возьмем какой-нибудь отрезок e и положим $e_n = \frac{1}{2^n}e$. Из свойства (в) (аддитивности) следует, что

$$2^n l(e_n) = l(2^n e_n) = l(e), \tag{3}$$

и то же для l'.

Пусть a — какой-либо отрезок. Откладывая на нем отрезки e_n, получим, по формуле (2),

$$m_n e_n \leqslant a < (m_n + 1)e_n.$$

Поэтому (пользуясь аддитивностью l) имеем

$$m_n l(e_n) \leqslant l(a) < (m_n + 1) l(e_n).$$

Отсюда, пользуясь (3) и деля на $l(e)$, получим

$$\frac{m}{2^n} \leqslant \frac{l(a)}{l(e)} < \frac{m_n + 1}{2^n}. \tag{4}$$

Применяя такой же вывод к функции l', получим

$$\frac{m_n}{2^n} \leqslant \frac{l'(a)}{l'(e)} < \frac{m_n + 1}{2^n}. \tag{5}$$

Эти неравенства (4), (5) верны при всяком n. Поэтому

$$\frac{l'(a)}{l'(e)} = \frac{l(a)}{l(e)}, \quad l'(a) = l(a)\frac{l'(e)}{l(e)}, \tag{6}$$

что и требовалось доказать. □

Единственность длины. Из доказанной теоремы непосредственно следует единственность длины при данном масштабе:

Теорема 3б. *При данном масштабе* e *существует только одна функция отрезка* $l(a)$ *с условиями* (а) — (г).

Доказательство. Если функции l, l' определены при одном масштабе e, т. е. $l(e) = l'(e) = 1$, то из (6) следует $l'(a) = l(a)$. □

Теорема 5 (о замене масштаба). *Если длины l, l' определены при масштабах e, e', то*

$$l'(a) = l(a)l'(e) = l(a) \cdot \frac{1}{l(e')}.$$

Доказательство. Если l при масштабе e, то $l(e) = 1$ и из (б) $l'(a) = l(a)l'(e)$. И совершенно так же, если $l'(e') = 1$, то $l(a) = l'(a)l(e')$. $\square$

Теорема 6 (о действиях с длинами). *Отношение равенства, операции суммы, разности, отношения «больше—меньше» для отрезков и их длин при любом данном масштабе полностью соответствуют друг другу; т. е. если $a = b$, то $l(a) = l(b)$, и обратно: если $l(a) = l(b)$, то $a = b$; и аналогично для сумм и т. д.*

Доказательство. Если $a = b$, то $l(a) = l(b)$ по условию (б). Докажем обратное. Пусть $l(a) = l(b)$. Допустим, однако, что $a \neq b$ и, скажем, $a > b$, т. е. $a = b + c$. Тогда $l(a) = l(b) + l(c) > l(b)$. Следовательно, необходимо $a = b$.

Если $a = b + c$, то $l(a) = l(b) + l(c)$ по условию (в). Докажем обратное. Пусть $l(a) = l(b) + l(c)$. По условию (в) $l(b) + l(c) = l(b + c)$, так что $l(a) = l(b + c)$. Отсюда, по только что доказанному, $a = b + c$. Для разности доказательство аналогично.

Если $a > b$, т. е. $a = b + c$, то $l(a) = l(b + c)$ и $l(a) = l(b) + l(c) > l(b)$. Докажем обратное. Пусть $l(a) > l(b)$. Тогда не может быть ни $a < b$, ни $a = b$, так как было бы $l(a) < l(b)$ или $l(a) = l(b)$. Следовательно, $a > b$. $\square$

Теорема 7 (об отрезке заданной длины). *При любом масштабе e для всякого положительного числа x существует отрезок с длиной (в масштабе e), равной x.*

Доказательство. Пусть задан масштаб e; длину в этом масштабе обозначаем l. Пусть дано еще число $x > 0$. Взяв целое $n > x$, построим отрезок $AB = ne$. На нем есть точки M, для которых $l(AM) < x$ (например, $AM = \frac{1}{2^m}e$, если $1/2^m < x$). Пусть F_1 — класс всех таких точек M, а F_2 — класс всех других точек N, лежащих на AB. Очевидно, эти классы удовлетворяют условиям аксиомы непрерывности. Поэтому существует такая точка C, что для всех $M \in F_1$ и $N \in F_2$

$$l(AM) \leqslant l(AC) \leqslant l(AN).$$

Отрезок AC имеет нужную длину: $l(AC) = x$, Действительно, допустим, $l(AC) < x$. Тогда возьмем отрезок b с $l(b) < x - l(AC)$. Отложив

его на CB, получим отрезок CD, и будет

$$l(AD) = l(AC) + l(CD) < l(AC) + x - l(AC) = x.$$

Тем самым $D \in F_1$, и потому должно быть $l(AD) \leqslant l(AC)$, хотя $l(AD) > l(AC)$. Это противоречие показывает, что не может быть $l(AC) < x$. Аналогично докажем, что не может быть $l(AC) > x$. Следовательно, $l(AC) = x$. Теорема 7 доказана. □

Теорема 7 вместе с теоремами 3 и 6 приводит к выводу:

Измерение отрезков, когда выбран масштаб, устанавливает взаимно однозначное соответствие между классами равных отрезков и положительными числами, при котором суммы отрезков и чисел соответствуют друг другу (*а вследствие этого и разности, и отношения «больше — меньше» также соответствуют друг другу*).

§ 4. Плоскостные аксиомы

Плоскостные аксиомы, обосновывающие измерение углов, можно сформулировать следующим образом, с наименьшим отклонением от аксиом, высказанных в гл. I ч. 2. Вслед за четырьмя группами линейных аксиом они образуют пятую группу.

V$_1$. *Существует три точки, не лежащие на одном отрезке.*

V$_2$. *Всякая прямая делит все не принадлежащие ей точки на два класса: «лежащие с одной и с другой стороны от нее».*

Замечание. В гл. I ч. 2 аксиома деления плоскости содержала еще утверждение, что всегда есть точки, лежащие с разных сторон от прямой. Но теперь введена аксиома V$_1$, из которой это непосредственно следует.

Действительно, если a — данная прямая, то по аксиоме V$_1$ существует точка A, не лежащая на ней. Возьмем на a точку B, проведем отрезок AB и продолжим за точку B. Его концы будут лежать с разных сторон от a.

Помимо этого аксиома V$_1$ обеспечивает вместе с аксиомами связи отрезков и точек, существование отрезков.

VI. Аксиомы угла.

Теперь, определив угол и поперечины, как в гл. I ч. 2, сформулируем аксиомы об углах.

VI$_1$ (аксиома равных углов). *Если у двух углов есть равные соответственные поперечины, то все их соответственные поперечины равны.* Такие углы мы называем *равными.*

VI$_2$ (аксиома откладывания угла). *От любого данного отрезка от данного его конца и по данную сторону от него можно отложить угол, равный данному неразвернутому углу.*

VI$_3$ (аксиома единственности). *Угол, отложенный по аксиоме* VI$_2$, *единственный.*

Замечание. Аксиома VI$_3$ на самом деле лишняя: она выводится из других. Аксиомы VI$_{1,\,2}$ можно заменить более простыми. Но пока мы принимаем все три аксиомы ради простоты.

VII. Аксиома параллельных.

В гл. I ч. 2 были даны три формулировки этой аксиомы; можно принять любую из них. А можно, например, заменить аксиому параллельных отрезков следующей.

Существует хотя бы один прямоугольник, или точнее: существуют такие четыре отрезка AB, AC, BD, CD, что $AC = BD$, $AB = CD$, $AC \perp AB$, $AB \perp BD$.

Другой вариант плоскостных аксиом:

В этом варианте аксиомы V и VI заменяются такими, которые говорят о треугольниках. Поэтому их можно объединить в одну группу:

V*. Аксиомы треугольника.

V$_1$*. *Существуют три точки, не лежащие на одном отрезке.*

Отрезки AB, BC, CA, соединяющие три точки A, B, C, образуют треугольник. Поэтому эту аксиому можно назвать «аксиомой существования треугольника».

V$_2$*. *Если отрезок пересекает сторону треугольника, то он* (*или его продолжение*) *пересекает еще одну сторону треугольника, либо проходит через вершину.* (Отрезок a пересекает отрезок b, если у отрезков a, b есть единственная общая точка, лежащая на них обоих. Аналогично, прямая *пересекает* отрезок, если она содержит одну, и только одну лежащую на нем точку; рис. 7.)

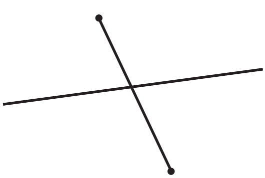


Рис. 7

Эта аксиома называется «аксиомой Паша» по имени немецкого геометра М. Паша (1843–1930), впервые высказавшего ее в 1882 г. Ее обычно выражают в следующей равносильной форме:

Если прямая пересекает сторону треугольника и не проходит через его вершину, то она пересекает и другую сторону.

Аксиома деления плоскости, как будет доказано, следует из аксиомы Паша. Но аксиома Паша гораздо проще, так как говорит не о

всей плоскости, а лишь о том или ином треугольнике, — собственно, о конечном числе точек.

Для формулировки дальнейших аксиом примем определение. Пусть a — данный отрезок и A — точка, не принадлежащая никакому отрезку, содержащему a. Мы говорим, что точка B лежит *с той же стороны* от отрезка a, что и точка A, если отрезок AB не имеет общих точек ни с каким отрезком, содержащим a. Соответственно, выражение «с данной стороны от отрезка a» означает «с одной стороны с данной точкой». (Как сказано, из аксиомы Паша следует, что есть только две «стороны» относительно любого данного отрезка, но для формулировки следующих аксиом это не нужно.)

V$_3$*. *Если $AB = A'B'$, $AC = A'C'$, $BC = B'C'$, и если D, D' — такие точки на AB и $A'B'$, что $AD = A'D'$, то также $CD = C'D'$* (рис. 8).

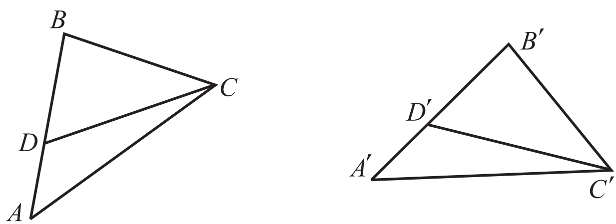


Рис. 8

V$_4$*. *Для любого треугольника ABC и отрезка $A'B'$, равного AB, с любой данной стороны от $A'B'$ существует такая точка C', что $A'C' = AC$, $B'C' = BC$* (рис. 9).

Аксиома V_3* заменяет аксиому VI_1 равенства углов, а аксиома V_4* — аксиому VI_2 откладывания угла; аксиома единственности VI_3 следует из них.

Теперь ее можно выразить в виде следующей теоремы:

Если точки C, C_1 лежат с одной стороны от отрезка AB и $AC = AC_1$, $BC = BC_1$, то эти точки совпадают.

Однако вывод аксиом VI из аксиом V_3*, V_4* мы излагать здесь не будем. Его можно найти в книге: *Александров А. Д.* Основания геометрии. М.: Наука, 1987. С. 150–153.

Вывод аксиомы деления плоскости из аксиомы Паша.

Теорема 1. *Если луч с началом в вершине угла пересекает одну поперечину, то он пересекает и всякую другую поперечину этого угла.*

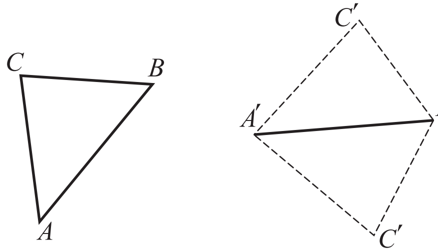


Рис. 9

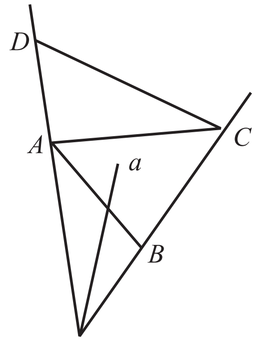


Рис. 10

Доказательство. Пусть луч a с началом в вершине угла пересекает поперечину AB (рис. 10). Тогда, по аксиоме Паша, он пересекает любую поперечину AC, так как сторону BC треугольника ABC он не может пересекать, поскольку содержит вершину угла. Точно так же луч пересекает любую поперечину DC, что и требовалось доказать. □

Теорема 2. *Прямая, пересекающая две стороны треугольника, не имеет с ним других общих точек.*

Доказательство. Пусть прямая a проходит через точки P, Q на сторонах AB, AC треугольника ABC. Других общих точек с этими сторонами прямая a не имеет (если бы была еще общая точка B_1, например, с AB, то прямая содержала бы AB).

Допустим, прямая a пересекает BC в точке D. Отрезок PQ, как и сторона BC, являются поперечинами угла A. Поэтому луч AD пересекал бы PQ в некоторой точке R. Получалось бы, что отрезок AD имеет две точки на прямой a и, следовательно, содержится в ней; прямая a проходила бы через A, что исключено. Теорема 2 доказана. □

Теорема 3. *Выполняется утверждение аксиомы деления плоскости* V_2.

Доказательство. Пусть a — данная прямая. По аксиоме V_1 есть точка, не лежащая на a. Соединим эту точку A с какой-либо точкой $P \in a$ и продолжим отрезок AP за точку P. Получаем отрезок AB, пересекающий прямую a.

Точки M такие, что AM не имеет с a общих точек, относим в класс F_A, остальные точки, не принадлежащие a, — в класс F_B; в этом классе точка B. Если M, $M' \in F_1$, то MM' не пересекается с a (так как иначе, по аксиоме Паша, a должна была бы пересекать AM или AM').

Пусть N, $N' \in F_2$; это значит, что AN и AN' пересекают a, а следовательно, по теореме 2, NN' не пересекает a. Таким образом, все точки, не принадлежащие прямой a, разделены на два класса, как и требуется аксиомой деления плоскости. □

§ 5. Алгебра углов. Измерение углов

Для углов имеет место «алгебра», аналогичная алгебре отрезков, с той, однако, разницей, что углы ограничены развернутым углом, а отрезки не ограничены.

Теорема 1. *Отношение равенства углов рефлексивно, симметрично и транзитивно.*

Доказательство. Равенство углов определяется равенством соответственных поперечин, т. е. через равенство отрезков. Поэтому оно так же рефлексивно, симметрично и транзитивно, как равенство отрезков (проследите это заключение в деталях). □

Дальнейшие выводы, относящиеся к алгебре углов, основаны на теоремах о равенстве треугольников. Эти теоремы доказаны в гл. II ч. 2 на основе аксиомы о равенствах углов, так что мы можем и теперь ими пользоваться, прежде всего в следующей теореме о смежных углах.

Теорема 2. *Углы, смежные равным углам, равны.*

Доказательство. Пусть O, O_1 — равные углы (обозначаемые их вершинами). На их сторонах отложим равные отрезки: $OA = O_1A_1$ — на общих сторонах этих углов и им смежных, и $OB = O_1B_1$ — на других сторонах. На продолжении этих сторон отложим также отрезки $OC = O_1C_1$ (рис. 11), так что $BC = B_1C_1$.

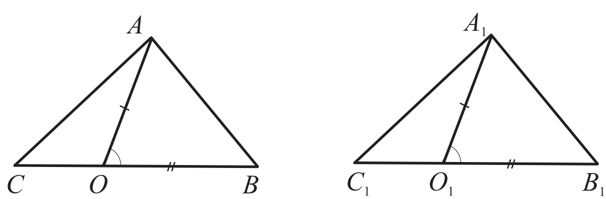


Рис. 11

Треугольники OAB, $O_1A_1B_1$ равны по двум сторонам и углу между ними (так как $\angle O = \angle O_1$ по условию). Поэтому $\angle B = \angle B_1$ и $AB = A_1B_1$. И треугольники ABC, $A_1B_1C_1$ равны по двум сторонам и углу между ними ($\angle B = \angle B_1$). Поэтому $AC = A_1C_1$.

Таким образом, у треугольников OAC, $O_1A_1C_1$ стороны равны, а

значит, равны их углы при вершинах O, O_1, что и требовалось доказать. □

Из теоремы 2 выводится обратная ей

Теорема 3. *Пусть углы* ab, a_1b_1 *равны, угол* bc — *смежный с* ab *и к углу* a_1b_1 *«пристроен» угол* b_1c_1, *равный* bc. *Тогда этот угол оказывается смежным с углом* a_1b_1 (рис. 12). («Пристроен» — значит, сторона b_1 у них общая, а a_1 и c_1 лежат по разные стороны от нее.)

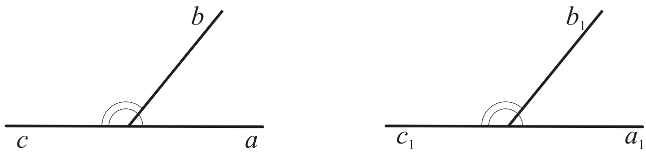


Рис. 12

Доказательство. Пусть выполнены высказанные условия. Продолжив сторону a_1 угла a_1b_1 за вершину, получаем угол b_1c_2, смежный с a_1b_1. И так как $\angle a_1b_1 = \angle ab$, то, по теореме 2, $\angle b_1c_2 = \angle bc$. Но, по аксиоме откладывания угла, от отрезка b_1 от данного его конца можно отложить по одну сторону от b_1 только один угол, равный данному $\angle bc$. Следовательно, $\angle b_1c_1$ совпадает с $\angle b_1c_2$, т. е. он смежный с a_1b_1, что и требовалось доказать. □

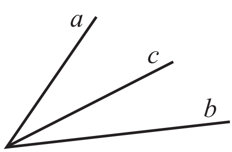


Рис. 13

Мы говорим, что угол ab *составлен* из углов ac, bc, если отрезок c проходит в угле ab; в этом случае мы также говорим, что $\angle ab$ *слагается* из $\angle ac$, $\angle bc$ или является их *суммой* (рис. 13).

Если угол ab развернутый, то *каждый* отрезок c с концом в его вершине (не налегающий на его стороны) условимся считать проходящим внутри угла ab и угол ab считать составленным из углов ac и bc. Углы эти, как было определено, смежны.

Для сложения углов выполняется теорема, аналогичная аксиоме сложения отрезков.

Теорема 4 (о сложении углов). *Углы, составленные из равных углов, равны, т. е. если отрезки* c, c_1 *проходят внутри углов* ab, a_1b_1 *и при этом* $\angle ac = \angle a_1c_1$, $\angle bc = \angle b_1c_1$, *то также* $\angle ab = \angle a_1b_1$ (рис. 14).

Доказательство. Проведем доказательство для случая, когда «сумма» — настоящий (не развернутый) угол. Пусть даны углы ab, a_1b_1 с вершинами O, O_1, причем угол ab — настоящий. Пусть отрезки c, c_1

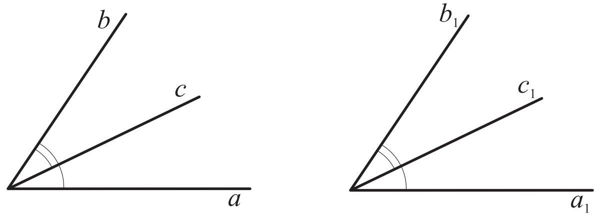


Рис. 14

проходят внутри углов ab и a_1b_1, причем $\angle ac = \angle a_1c_1$, $\angle bc = \angle b_1c_1$. Так как угол ab — настоящий, то отрезок c пересекает какую-то его поперечину AB (A на a, B на b). Пусть C — точка пересечения отрезка c и поперечины AB (рис. 15). На отрезках a_1, b_1, c_1 возьмем точки A_1, B_1, C_1 так, что $O_1A_1 = OA$, $O_1B_1 = OB$, $O_1C_1 = OC$ (если отрезки a_1, b_1, c_1 «не дотягивают» до этого, то их можно продолжить). Тогда по равенству углов при O и O_1 будет

$$\triangle OAC = \triangle O_1A_1C_1, \quad \triangle OBC = \triangle O_1B_1C_1. \tag{1}$$

Тем самым углы при C и C_1 в этих треугольниках равны.

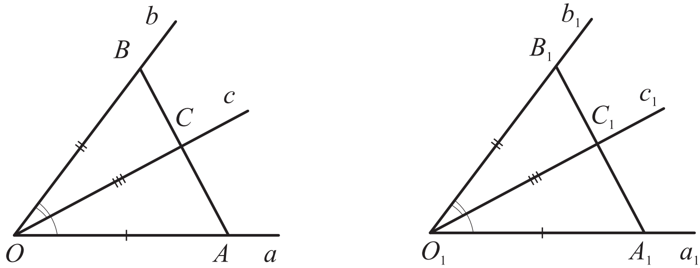


Рис. 15

Но углы при C — смежные. Поэтому углы при C_1 тоже смежные, как следует из теоремы 3. Стало быть, отрезки A_1C_1, B_1C_1 образуют один отрезок A_1B_1. И так как, ввиду (1), $A_1C_1 = AC$, $B_1C_1 = BC$, то, по аксиоме сложения, также $A_1B_1 = AB$. Таким образом, у треугольников OAB, $O_1A_1B_1$ все стороны равны, а значит, $\angle O = \angle O_1$, т. е. $\angle ab = \angle a_1b_1$, что и требовалось доказать. $\square$

Теорема 4 сложения углов доказана, когда хотя бы один составной угол — настоящий. Тогда выходит, что и другой — настоящий. Случай, когда оба угла развернутые, решается следующей теоремой.

Теорема 5. *Все развернутые углы равны.*

Доказательство. На сторонах двух развернутых углов O, O_1 возьмем точки A, B и A_1, B_1 так, что $OA = O_1A_1$, $OB = O_1B_1$. Тогда, по теореме 1 § 6 гл. I ч. 2 о расположении отрезков, отрезки OA, OB и O_1A_1, O_1B_1 составляют отрезки AB, A_1B_1 (рис. 16). Поэтому, согласно аксиоме сложения отрезков, $AB = A_1B_1$, т. е. поперечины равны, и, стало быть, $\angle O = \angle O_1$, что и требовалось доказать. □

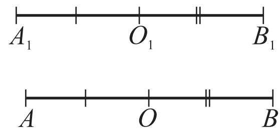


Рис. 16

Теперь мы обобщим понятие о сложении углов буквально так же, как мы обобщили в начале § 2 понятие о сложении отрезков.

Определение. Мы говорим, что угол α представляет *сумму* углов β и γ, и пишем $\alpha = \beta + \gamma$, если угол α составлен из углов, равных β и γ, или, вообще, равен углу, составленному из таких углов.

Операцию сложения данных углов β и γ можно представить так. Строим угол ac, равный β, и пристраиваем к нему угол cb, равный γ, т. е. строим угол, равный γ, со стороной c — той же, что у угла ac, но с другой стороны от нее (рис. 17). При этом может оказаться, что сторона c уже не будет проходить между a и b: суммарный угол получится больше развернутого.

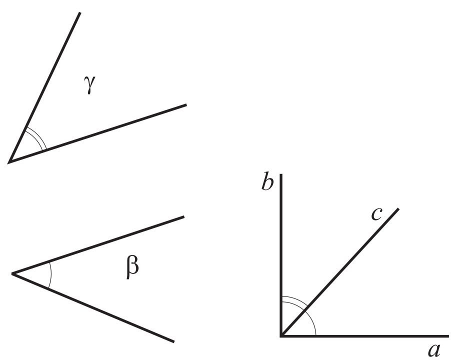


Рис. 17

Дальше мы определим и рассмотрим такие углы, но пока мы этого делать не будем, ограничиваясь сложением углов, когда сумма оказывается, самое большее, развернутым углом.

Так же, как для отрезков, определяем при натуральном n угол $n\alpha$ и угол $\frac{1}{n}\alpha$ (помня, что угол $n\alpha$ должен быть не больше развернутого).

Далее, так же, как для отрезков, определяем вычитание углов: называем их *разностью* $\gamma = \alpha - \beta$ такой угол, что $\beta + \gamma = \alpha$, если есть такой γ. Точно так же определяем: $\alpha > \beta$, если есть такой угол γ, что $\alpha = \beta + \gamma$ (рис. 17).

Как и для отрезков, доказывается

Теорема 6. *Если $\alpha > \beta$ и $\alpha_1 = \alpha$, $\beta_1 = \beta$, то $\alpha_1 > \beta_1$, $\alpha_1 - \beta_1 = \alpha - \beta$.*

Эта теорема отличается от соответствующей леммы из § 2 для отрезков лишь тем, что вместо концов A, B и подобных отрезков появляются стороны углов a, b и т. п. (большие буквы заменяются на строчные) и вместо откладывания отрезка вдоль данного отрезка появляется откладывание угла с той же стороны, где данный угол.

Соответственно и доказательство получается такой заменой в доказательстве для отрезков, а вместо ссылки на аксиому сложения отрезков ссылаемся на доказанную теорему о сложении углов.

Читателю полезно самому провести получающееся доказательство. □

В итоге получаем:

Теорема 7 (об алгебре углов). *Для углов наряду с основным понятием равенства определены операции сложения и вычитания и отношение «больше—меньше». При этом оказывается, что эти операции и отношение обладают такими же свойствами, как для положительных чисел, не превосходящих какого-либо данного числа* (*соответственно тому, что углы не больше развернутого*). *Именно, прежде всего выполняется следующее:*

1. *Для любых двух углов α, β таких, что $\beta \leqslant \beta'$, где β' — угол, смежный с α, определена их сумма.*

2. *Выполняются переместительный и сочетательный законы сложения.* □

Выполняются также пять свойств, которые для отрезков сформулированы в конце § 2.

Доказательства всех этих свойств получаются точно так же. Проследить все эти выводы мы предоставляем читателю в качестве очень полезного упражнения, когда доказательство не просто повторяется, а повторяется с некоторыми небольшими изменениями в других условиях.

В частности, при натуральном n, если $\alpha = \beta$, то $\frac{1}{n}\alpha = \frac{1}{n}\beta$. Это включает, что *половины равных углов равны* (и так как угол равен са-

мому себе, то его половины равны; соответственно, биссектриса угла только одна). Применяя это к прямому углу, т. е. к половине развернутого угла, получаем:

(1) *Все прямые углы равны.*

(2) *Через данную точку на отрезке проходит только один перпендикуляр* (*один с точностью да удлинения и укорочения*).

Сверхтупые углы. Прибавляя какой-нибудь угол к развернутому углу, получим «сверхтупой угол». Два неналегающих отрезка с общим концом определяют два угла, либо развернутые, либо если один не развернутый, то другой сверхтупой. Он отличается тем, что если продолжить сторону угла за вершину, то он представится как сумма развернутого и настоящего угла. Какую из сторон мы продолжим — не важно: углы, прибавляемые в каждом случае к развернутому, — вертикальные, а потому равны. Вообще, определяем два сверхтупых угла как равные, если они получаются прибавлением к развернутым углам равных углов. Мы получаем возможность складывать углы и тогда, когда в сумме получается сверхтупой угол. При этом аналогично теореме 4 выполняется

Теорема 8. *Если равные углы* $\alpha = \alpha_1$, $\beta = \beta_1$ *образуют в сумме сверхтупые углы* $\gamma = \alpha + \beta$, $\gamma_1 = \alpha_1 + \beta_1$, то $\gamma = \gamma_1$.

Доказательство читатель получит сам, пользуясь определением равенства сверхтупых углов и свойствами сложения и вычитания обычных углов. □

Сверхтупой угол, как и обычный угол, можно еще представлять как «часть плоскости» — как фигуру, образованную отрезками (лучами) с общим концом (когда говорят о том, чтобы вырезать угол из бумаги, угол понимают в этом смысле). Сложение углов заключается тогда в том, что сумма составляется из слагаемых. Угол γ является суммой углов α, β, если он служит их объединением и они имеют только общую сторону.

Стороны настоящего угла делят все отрезки, исходящие из его вершины (не считая самих сторон), на два класса: те, которые проходят в угле, т. е. пересекают его поперечину, и все остальные. Каждый из этих классов вместе со сторонами угла образует «плоский угол»: один меньше развернутого, содержащийся в полуплоскости, другой больше развернутого, содержащий полуплоскость, — сверхтупой.

Термин «плоский угол» выражает здесь то, что речь идет о «части плоскости».

Измерение углов. Для углов выполняется утверждение, совершенно аналогичное аксиоме непрерывности.

Теорема 9. *Пусть отрезки, проходящие в данном (настоящем) угле ab, разбиты на два класса F_1, F_2 так, что если $d \in F_1$ и $e \in F_2$, то угол ad не больше угла ae. Тогда существует такой идущий внутри угла ab отрезок c, что для всех $d \in F_1$ и $e \in F_2$ $\angle ad \leqslant \angle ac \leqslant \angle ae$.*

Доказательство. Проведем в угле ab поперечину AB. Отрезки, проходящие в угле, пересекают ее, и этим осуществляется взаимно однозначное соответствие между отрезками, идущими в угле, и точками поперечины. При этом соотношения углов ad и отрезков $AD \subset AB$ соответствуют друг другу. Поэтому сказанное в теореме 9 автоматически следует из аксиомы непрерывности. □

Доказанное также автоматически приводит к тому, что измерение углов осуществляется подобно измерению отрезков. Достаточно повторить выводы § 3, заменяя отрезки углами.

Так обосновывается измерение настоящих углов (поскольку лишь такие углы фигурируют в доказанной теореме). Но отсюда измерение развернутого и сверхтупых углов получаются тем, что мы представляем их как суммы настоящих углов.

За единицу измерения углов принимают градус: $\frac{1}{90}$ прямого угла. Это чистая условность, и включать ее в аксиомы, как порой делают, нелепо, как будто геометрия зависит от единицы измерения углов.

Задача. Сформулировать для углов теоремы, аналогичные теоремам об отрезках: о существовании половины из § 2 и, далее, об измерении — все теоремы § 3. Приведите их доказательства. (Не забудьте о том, что углы ограничены.)

§ 6. Пространственные аксиомы

Пространственные аксиомы были изложены в § 8 гл. I ч. 2 и затем для пространства любого числа измерений — в § 1 гл. V ч. 3. Здесь мы повторим, а затем дополним это изложение.

В пространственной аксиоматике наряду с основными понятиями планиметрии вводятся новые объекты: *плоскости* и отношение — точка *принадлежит* плоскости, или, что то же, плоскость *проходит через* точку. Вводится определение: отрезок *содержится* (*лежит*) в плоскости, если все его точки принадлежат этой плоскости; так же определяется, что прямая *содержится* в плоскости.

Аксиомы делятся на: 1) линейные, 2) плоскостные, 3) пространственные.

Линейные аксиомы дословно повторяют аксиомы § 1.

Плоскостные аксиомы повторяют аксиомы § 4 с той разницей, что теперь к трем из них (V_2, VI_2, VI_3) добавляется указание, что она относится к каждой плоскости.

Пространственные аксиомы. Их мы делим здесь на две группы: 1) аксиомы плоскости и 2) аксиомы размерности.

Аксиомы плоскости.

VIII$_1$. *Через каждые три точки проходит плоскость.*

VIII$_2$. *Если две плоскости имеют две общие точки, то их пересечение есть прямая (т. е. в них обеих содержится прямая, которая содержит все их общие точки и никакие другие).*

Аксиомы размерности.

IX. *Существует n и не более взаимно перпендикулярных прямых.*

Для трехмерного пространства аксиомы размерности можно формулировать так:

IX$_1$. *Существуют 4 точки, не лежащие в одной плоскости.*

IX$_2$. *Если две плоскости имеют общую точку, то они имеют по крайней мере еще одну общую точку.* Другими словами: *две плоскости не могут иметь только одну общую точку.*

Первая из этих аксиом означает, что пространство не сводится к плоскости и, стало быть, не менее чем трехмерно. Вторая означает, что пространство не более чем трехмерно.

Аксиома IX тоже содержит два утверждения.

IX. 1. *Существует n взаимно перпендикулярных прямых.*

IX. 2. *Существует не более n взаимно перпендикулярных прямых.* (Согласно определению равенства углов и прямого угла, две прямые a, b взаимно перпендикулярны, если у них есть единственная общая точка O и на a точки A, A_1 и на b точка $B \neq O$ такие, что $OA = OA_1$, $BA = BA_1$.)

В аксиомах стереометрии (гл. I, ч. 2) вместо двух аксиом VIII_2 и IX_2 была одна:

Пр. 3. *Если две плоскости имеют общую точку, то их пересечение есть прямая.*

Это очевидно следует из аксиом VIII_2 и IX_2.

Так же очевидно аксиомы VIII_2 и IX_2 следуют из Пр. 3.

Таким образом, аксиомы трехмерного пространства, изложенные здесь, и аксиомы стереометрии, изложенные раньше, равносильны.

Дополнение. Вводя в качестве основных объектов плоскости, мы отходим от той установки, что основные понятия должны как можно ближе соответствовать практике. Плоскость во всем ее бесконечном протяжении не только не имеет прямого прообраза в практике, но, собственно говоря, и не доступна прямому наглядному представлению. Можно, однако, дать аксиоматику пространства, в которой основные объекты только те же, что в планиметрии, а плоскость определяется через них. Это можно сделать следующим образом.

Определение. Пусть A, B, C — три точки, не лежащие на одном отрезке, так что имеется треугольник ABC. Будем говорить, что точка M *принадлежит плоскости* (ABC), если существует содержащий ее отрезок, имеющий с треугольником ABC две общие точки.

Плоскостью (ABC) называем фигуру, образованную указанными точками M; другими словами, плоскость — это геометрическое место точек с указанным свойством.

Поскольку плоскость не включается в число основных объектов, аксиомы надо формулировать без нее. В такой аксиоматике плоскостные аксиомы являются одновременно аксиомами стереометрии. Линейные аксиомы те же, что и раньше, но плоскостные аксиомы изменяются. Определяем понятие угла так же, как и раньше. Дальше имеется в виду настоящий (не развернутый) угол. Вводим следующие аксиомы.

V. Аксиомы угла.

$\mathbf{V}_1$. *Существуют три точки, не принадлежащие одному отрезку* (эта аксиома гарантирует существование настоящего угла).

$\mathbf{V}_2$. *Если отрезок, проведенный из вершины угла, пересекает какую-либо поперечину, то он или его продолжение пересекают любую другую.* Мы говорим, что отрезок проходит внутри данного угла.

$\mathbf{V}_3$. *Если у двух углов есть равные соответственные поперечины, то все их соответственные поперечины равны.*

$\mathbf{V}_4$. *Какие два угла ab, cd ни заданы, существует такой угол ae, равный cd, что либо его сторона e проходит в угле ab, либо b проходит в угле ae, либо e совпадает с b.* Угол откладывается, так сказать, в пространстве.

Аксиома параллельных отрезков. *Если отрезки AC, BD равны и перпендикулярны AB, и AD, BC пересекаются, то $CD = AB$* (*оговорка о пересечении AD и BC заменит условие, что отрезки AC, BD лежат в плоскости*).

Аксиома размерности. *Существуют три и не больше взаимно перпендикулярных отрезка, т. е. три отрезка, пересекающихся в одной точке и образующих прямые углы.*

Сформулированные аксиомы позволяют доказать, что если плоскость понимается в смысле данного выше определения, то для плоскостей выполняются все аксиомы стереометрии, как плоскостные, так и пространственные. Однако доказательство этого не просто. И это естественно: чем меньше мы требуем в аксиомах, тем более длинный путь выводов надо пройти, чтобы прийти к нужным результатам.

§ 7. Понятие фигуры

Аксиоматическое определение понятия фигуры уже было изложено в § 5 гл. I ч. 2. Здесь мы представляем его несколько иначе.

Аксиоматика фигуры.

Основные объекты: 1) *точки*, 2) *фигуры*. Основное отношение: точка *принадлежит* фигуре; в обозначении $A \in F$ и т. п.

Аксиомы.

1. *Фигура определяется своими точками, т. е. если имеются фигуры F_1 и F_2 такие, что каждая точка, принадлежащая одной из них, принадлежит также другой и обратно, то F_1* и *F_2 — одна и та же фигура.*

2. *Точка есть фигура; она принадлежит себе, и никакие другие точки ей не принадлежат.*

3. *Для всякого условия, налагаемого на точки, существует фигура, содержащая все точки с данным условием и никакие другие.* При этом имеется в виду условие, которое: 1) выражается через понятия, фигурирующие в принятой аксиоматике геометрии, и 2) является проверяемым для каждой точки — выполняется оно для нее или нет. (Эта аксиома фигурировала в § 5 гл. I ч. 2 под номером 5).

При нашей аксиоматике, например, то условие, что «точка принадлежит данному отрезку», считается проверяемым.

Смысл этих аксиом был разъяснен при их изложении выше (§ 5 гл. I ч. 2). Аксиома 3 выражает по существу то же, что и определение геометрического места. Однако, как было там же отмечено, в ней содержится некоторая неопределенность: какие условия считаются «в принципе проверяемыми». Уточним это, сформулировав аксиому 3 для элементарной геометрии с нашей аксиоматикой.

Условие, налагаемое на точки, назовем *элементарным*, если оно может быть выражено в понятиях аксиоматики так, что его провер-

ка для каждой точки проходит в конечное число шагов: при каждом шаге либо устанавливается одно из основных отношений, либо производится «построение» из тех, какие допущены аксиомами. Такими построениями являются проведение отрезка, откладывание отрезка или угла, равного данному (это мысленное построение есть не что иное, как фиксация существования указываемого объекта).

Аксиома 3′ (аксиома геометрического места в элементарной геометрии). *Для всякого элементарного условия, налагаемого на точки, существует фигура, содержащая все точки с данным условием и никакие другие.*

В предыдущем изложении аксиом фигуры (§ 5 гл. I ч. 2) среди них были 1) «аксиома отрезка» о том, что отрезок есть фигура, 2) аксиомы об операциях с фигурами, как их объединение и др. Теперь эти аксиомы представляются иначе — не как аксиомы.

То условие, что точка M принадлежит данному отрезку AB, считается проверяемым, оно, стало быть, элементарно. Поэтому согласно аксиоме 3′ точки отрезка AB образуют фигуру, обозначим ее $\overline{AB}$. Вместе с тем отрезок AB определяется своими точками (теорема 3 § 1). Это позволяет отождествить фигуру $\overline{AB}$ с самим отрезком AB и сказать: «отрезок есть фигура»[32].

Объединение и пересечение конечного числа фигур представляет собой фигуру, потому что условие, проверяемое в конечное число шагов для каждой из данных фигур, очевидно, проверяемо для них всех так же в конечное число шагов. Если число фигур бесконечно, то для них всех проверяемость условия само собой не обеспечена и требует особой проверки. Но, например, круг можно определить как объединение всех отрезков (на плоскости) с общим концом O, равных данному — радиусу. Принадлежит ли точка M данному кругу, проверяется очевидно: проводим отрезок OM и вдоль него откладываем радиус.

Дополнение фигуры, очевидно, есть фигура, поскольку при определении фигуры проверяется, выполняется для точки условие или нет; если нет, то точка и принадлежит дополнению.

В понятии элементарного условия можно иметь в виду не обязательно принятую нами аксиоматику, но и любую другую, которую можно назвать чисто геометрической, т. е. такую, в которой не фигурируют ни действительные числа, ни величины... Например, если основным объектом считается прямая, то «построение» может состоять в «проведении» прямой через две точки.

---

[32]Совершенно аналогично можно сказать, что «плоскость есть фигура».

Рассмотрим примеры, связанные с нашей аксиоматикой. Окружность, круг, прямая, луч, полуплоскость определяются элементарными условиями. Эллипс с фокусами F_1, F_2 можно определить как геометрическое место таких точек M, что из отрезков, равных F_1M, F_2M составляется данный отрезок AB. Это условие очевидно проверяемо: достаточно отложить отрезки, равные F_1M, F_2M, вдоль отрезка AB от точек A, B.

Эллипс можно также определить условием, что отношение длины отрезка F_1M к расстоянию точки M до директрисы равно данному числу. Это условие не элементарно, так как требует измерения длины, которое может не осуществиться в конечное число шагов. Но то же условие можно сформулировать иначе: отношение отрезка F_1M к перпендикуляру MN, опущенному на директрису, равно отношению данных отрезков. Такое условие проверяемо построением.

В общем, аксиома $3'$, как она сформулирована, выражает понятие фигуры в элементарной геометрии и позволяет дать следующее определение.

Элементарной геометрией называется теория, предмет которой составляют фигуры, определяемые элементарными условиями (как они были только что определены).

Можно, конечно, понимать «проверяемое условие» в более общем смысле, когда оно не будет элементарным, не проверяется в конечное число шагов, как, например, условие, требующее предельного перехода. Мы могли при прежде принятой аксиоматике с понятием численной длины считать проверяемым то условие, что отрезок имеет данную численную длину при данном масштабе.

Вообще, понятие проверяемого условия, а значит и понятие фигуры, допускает разные градации. Но первая и основная из них та, которая установлена определением элементарного условия и соответствующим ему понятием фигуры.

§ 8. Величина

Понятие величины является основным в точном естествознании; подавляющее большинство законов физики говорит о зависимостях между теми или иными величинами. Простейшая из величин — это длина; из практики она вошла в геометрию. Как понятно каждому, длина обладает следующими свойствами, которые мы сформулируем на интуитивном уровне без строгости.

1. Длины можно складывать (длины складываются, когда один отрезок прикладывается к другому).

2. Если к данной длине прибавляется еще длина, то получается бо́льшая длина.

3. Результат сложения — сумма длин — не зависит ни от порядка сложения, ни от того как объединяются слагаемые (коммутативность и ассоциативность) .

4. Длина может непрерывно изменяться.

Эти свойства длины, если их выразить точно в общем виде как аксиомы, дают аксиоматическое определение общего понятия величины (или, уточняя, — аддитивной, положительной скалярной величины; «аддитивной» потому, что для нее определено сложение[33], «скалярной» — чтобы отличить от векторных величин).

Аксиоматическое определение величины. *Величиной* называется элемент множества (совокупности) «однородных величин», в котором определена операция, называемая *сложением*, и выполняются следующие ее свойства (операция обозначается знаком $+$).

I. Аксиомы сложения.

I$_1$. *Для каждых двух величин a, b существует такая однозначно определенная величина c, что $c = a + b$ (иначе говоря, операция сложения сопоставляет каждой паре величин a, b определенную величину c* — «их сумму»).

I$_2$. $a + b = b + a$.

I$_3$. $(a + b) + c = a + (b + c)$.

II. Аксиома неравенства.

II. *Для каждых двух величин a, b верно одно из трех: либо* 1) $a = b$; *либо* 2) *существует такая однозначно определенная величина c, что $a = b + c$, либо наоборот*: 3) *существует такая величина d, что $b = a + d$.*

В случае 2) говорят, что *a больше b*: $a > b$; в случае 3) — что $b > a$. (Это вполне соответствует обыденному понятию: больше та величина, которая получается, когда к данной что-то прибавляют.)

Когда $a = b + c$, то полагают также $c = a - b$, т. е. если $a > b$, то существует однозначно определенная величина — *разность* $c = a - b$.

III. Аксиомы непрерывности.

III$_1$. *Для всякой величины есть меньшая.*

[33]Есть величины, для которых сложение не определено, например температура. Прибавляют не температуру, а количество тепла. Заряд — величина, которая бывает положительной и отрицательной.

III$_2$. *Всякая ограниченная сверху последовательность величин имеет точную верхнюю границу, т. е. если для последовательности величин* $a_1, a_2, a_3, \ldots$ есть такая c, что все $a_n \leqslant c$, то либо среди величин a_n есть наибольшая, либо существует такая величина a, что все $a_n < a$ и при всякой величине b найдется такое n, что $a < a_n + b$ (или, что то же, $a - a_n < b$).

Замечание. Если аксиому III$_1$ заменить на противоположную, — ту, что среди величин существует наименьшая, — то получим определение *дискретной величины*, измеряемой натуральными числами, как численность совокупности предметов.

Выполняется следующая важнейшая теорема.

Теорема 1 (об измерении величин). *Величинам (любого данного типа) можно взаимно однозначно сопоставить положительные числа так, что суммам величин будут отвечать суммы чисел. Такое соответствие однозначно определяется выбором той величины* e, *которой сопоставляется число* 1.

Такое сопоставление чисел величинам называется *измерением*, а величина e — *единицей измерения* или *масштабом.*

Доказательство этой теоремы по существу аналогично доказательству теорем о численной длине в § 3. □

В планиметрии есть три величины: длина, угол (как величина) и площадь. Им можно дать следующие определения, совершенно сходные друг с другом.

Длиной отрезка называется величина, отнесенная отрезкам так, что выполнены два условия:

1. У равных отрезков длина одна и та же.

2. Если отрезок a составлен из отрезков a_1 и a_2, то его длина равна сумме их длин.

Площадью фигуры, составленной из многоугольников, называется величина, относимая таким фигурам, с двумя условиями:

1. У равных фигур площадь одна и та же.

2. Если фигура F составлена из двух фигур F_1, F_2 (т. е. служит их объединением, но эти фигуры не имеют общих внутренних точек), то площадь фигуры F равна сумме площадей фигур F_1 , F_2. (Определение площади для немногоугольных фигур сложнее и будет дано в гл. II.)

Величиной угла называется величина, относимая углам так, что выполнены два условия:

1. У равных углов величина одна и та же.

2. Если угол α составлен из углов α_1, α_2, то его величина равна сумме их величин.

Следует оговорить, что величина угла — это величина «с ограничением»: она не может быть больше величины развернутого угла или, при другом взгляде, больше двух развернутых углов. Поэтому здесь аксиомы сложения должны быть дополнены: существует такая величина d, что сумма $a + b$ определена лишь тогда, когда $a \leqslant 2d - b$, где $2d$ — величина развернутого угла.

Величины всегда относятся к каким-либо объектам, как «величина объекта данного класса», как длина, относящаяся к отрезкам (или вообще — к спрямляемым кривым), площадь к многоугольникам (или более общим фигурам с внутренностью), как в обыденной жизни и в физике масса или вес относится к телам и т. п. Величина — это свойство объекта данного класса, которое в каком-то отношении может быть больше или меньше и притом так, что позволяет точное сравнение, называемое измерением. Поэтому говорят, что величина — это то, что можно измерить. Сложение же определяется не для любых величин. Но в математике имеют в виду только аддитивные величины.

Глава II

ПЛОЩАДЬ И ОБЪЕМ

§ 1. Определение площади

Площадь многоугольных фигур. Мы будем рассматривать только ограниченные фигуры, и слово «фигура» будет всегда обозначать плоскую ограниченную фигуру, т. е. такую, которая содержится в каком-нибудь круге.

Будем говорить, что фигура составлена из нескольких фигур, если она служит их объединением и никакие две из них не имеют общих внутренних точек. Под многоугольной фигурой будем всюду ниже понимать *многоугольную площадку* (см. § 4 гл. II ч. 2); рис. 18. Проще говоря, многоугольная фигура — это объединение конечного числа треугольников.

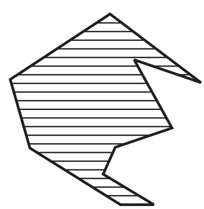


Рис. 18

Площадью многоугольной фигуры называется величина, обладающая следующими двумя свойствами:

1) *равные фигуры имеют равные площади*;

2) *если фигура составлена из нескольких многоугольных фигур, то ее площадь равна сумме площадей этих фигур.*

Первое свойство — инвариантность, неизменяемость при перемещениях, второе — аддитивность; его достаточно требовать для того случая, когда фигура слагается из двух фигур.

Удобно ввести следующие обозначения.

Равенство фигур обозначим так: $F \cong F'$. То, что фигура F составлена из фигур F_1, F_2 и т. д., будем обозначать как сумму: $F = F_1 + F_2$. Площадь будем обозначать буквой S.

Вместо самой площади как величины удобно рассматривать ее численное значение при данной единице измерения и дать следующее определение.

«Численной площадью» — численным значением площади фигуры F при данной «единичной» фигуре E — называется число $S(F)$, относимое многоугольной фигуре, каждой — свое, так что выполнены условия:

1) $S(F) > 0$,
2) если $F \cong F'$, то $S(F) = S(F')$,
3) $S(F_1 + F_2) = S(F_1) + S(F_2)$,
4) $S(E) = 1$.

Последнее условие означает, что единичная фигура — это та многоугольная фигура E, которой отнесено численное значение площади, равное единице. В качестве такой фигуры берут «единичный квадрат», т. е. квадрат со стороной, равной выбранной единице длины. Но это совершенно не обязательно: единичной фигурой может быть, в принципе, любая многоугольная фигура.

Заметим, что из аддитивности площади (свойства 2) следует: *если фигура F содержит F_1 и не совпадает с F_1, то $S(F) > S(F_1)$.*

Действительно, если $F \supset F_1$ и $F \neq F_1$, то очевидно $F = F_1 + F_2$, где F_2 — тоже многоугольная фигура (рис. 19).

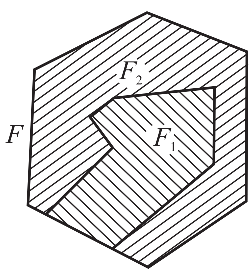


Рис. 19

По аддитивности $S(F) = S(F_1) + S(F_2)$, и поэтому
$$S(F) > S(F_1).$$

Дальше мы будем говорить о площади, допуская, что под «площадью» можно разуметь как саму величину, так и ее численное значение при какой-нибудь единице измерения, которую можно для дальнейшего выбрать раз навсегда. В большинстве случаев

удобнее иметь в виду именно численное значение, когда говорится о сложении и сравнении площадей (но не об их измерении: измеряется величина, а ее численное значение — это результат измерения).

То, что у каждой многоугольной фигуры, в частности у квадрата, есть определенная площадь, представляется, в общем, как нечто само собой разумеющееся. Но можно задать следующий вопрос.

Представим себе, что, разбив квадрат или другой многоугольник на какие-то многоугольники, мы перемещаем их так, чтобы они не налегали друг на друга. Мы будем получать различные новые фигуры. Не может ли при этом получиться такая фигура, которая уместится внутри первоначального многоугольника или внутри одной из полученных из него новых фигур? Если бы это случилось, то мы имели бы две фигуры, у которых, с одной стороны, площади должны быть равны, так как они составлены из попарно равных фигур. С другой стороны, у фигуры, умещающейся внутри другой, площадь меньше. То есть получалось бы противоречие. Выходило бы, что понятие о равенстве или неравенстве площадей и, стало быть, само понятие о площади оказались лишенными смысла.

Мы скажем: «не может быть, чтобы одна фигура уместилась внутри другой, ведь у них площадь одна и та же». Но именно об этом и стоит вопрос: имеет ли смысл понятие площади? Иначе говоря: существует ли в самом деле такая величина, как площадь, у многоугольных фигур?

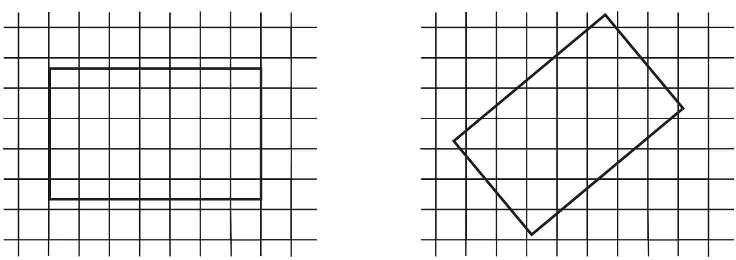


Рис. 20

Можно поставить и другой вопрос. Площадь прямоугольника определяют, покрывая его квадратами (рис. 20, *а*), и квадраты эти берутся такими, что стороны их параллельны сторонам прямоугольника. А что будет, если брать квадраты, повернутые относительно прямоугольника, или, что равносильно, повернуть прямоугольник, как на рис. 20, *б*? Будет ли подсчет таких квадратов давать то же самое (в пределе,

когда квадраты уменьшаются)? Иначе говоря, будет ли (в пределе) площадь маленьких квадратов, укладывающихся на прямоугольнике, той же самой, с точностью до квадратов, которые пересекаются со сторонами прямоугольника?

Мы уверены, что результат будет тот же, потому что у фигуры есть определенная площадь, так что ее измерение должно давать всегда один и тот же результат. Но именно об этом и идет речь: существует ли у фигуры определенная площадь, независимая от того, как мы ее измеряем?

Приведенные рассуждения показывают, что существование площади, т. е. существование величины со свойствами 1, 2, вовсе не так очевидно, как кажется на первый взгляд. Существование ее нужно доказать[34].

Доказательство может быть дано. Именно, доказывается следующая теорема.

Теорема I. *Каждая многоугольная фигура имеет определенную площадь.*

Для численных значений это можно выразить так.

При заданной единичной фигуре E каждой многоугольной фигуре отвечает, и притом единственная, численная площадь со свойствами 1, 2, 3, 4.

Если фигура E заменяется другой E', то все численные площади изменяются на один и тот же множитель:

$$S'(F) = kS(F), \quad k = S'(E) = \frac{1}{S(E')}.$$

Ту же теорему можно высказать иначе.

Существует, и притом единственная, функция со свойствами 1–4, *определенная на множестве многоугольных фигур.* Функции, соответствующие разным «единичным» фигурам E, отличаются множителем.

Площади общих фигур. Вслед за определением площади многоугольных фигур встает вопрос об определении площади для других фигур. Ее определяют, обобщая тот способ, каким в школьном курсе находят площадь круга.

---

[34]На этот вопрос обратил внимание итальянский математик Де-Цольт (в 1881 г.) и пытался решить его, но только потом это было строго сделано профессором университета в Одессе С. О. Шатуновским, Д. Гильбертом и др.

Пусть F — какая угодно данная фигура. Будем рассматривать многоугольные фигуры — содержащие F и содержащиеся в F; первые обозначим G и их площади — $S(G)$; вторые обозначим H и их площади $S(H)$ (рис. 21). (Если многоугольных фигур H, содержащихся в данной фигуре F, нет, то считаем $S(H) = 0$).

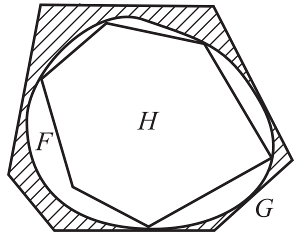


Рис. 21

Если многоугольная фигура G содержит F, а $F \supset H$, то, очевидно, $G \supset H$, и, значит, $S(G) \geqslant S(H)$ (при этом $S(G) = S(H)$ возможно только в том случае, когда обе фигуры G и H совпадают и, значит совпадают с F, так что сама фигура F — многоугольная).

В качестве площади фигуры F можно взять такую величину $S(F)$, что для всех фигур G и H

$$S(G) \geqslant S(F) \geqslant S(H).$$

То есть за площадь фигуры F принимается величина, которая не больше площадей многоугольных фигур, содержащих F, и не меньше площадей многоугольных фигур, содержащихся в F.

Однако площади фигур G и H могут отличаться так, что между ними будет целый интервал величин, и, стало быть, величина $S(F)$ не будет единственной. Площадь оказывается неопределенной. Она будет определенной, если разность площадей $S(G)$, $S(H)$ может быть сколь угодно малой.

Таким образом, мы приходим к окончательному определению, которое, как легко видеть, воспроизводит в общем виде школьное определение площади круга.

Площадью фигуры F называется величина, которая не больше площадей многоугольных фигур, содержащих F, и не меньше площадей многоугольных фигур, содержащихся в F, при условии, что разности этих площадей могут быть сколь угодно малыми.

Мы говорим тогда, что *фигура F имеет определенную площадь.* Всякая многоугольная фигура F попадает под это определение, поскольку тогда сама F оказывается многоугольной фигурой, содержащей F и содержащейся в F.

Оказывается, что так определенная площадь и для немногоугольных фигур обладает теми же свойствами, какие определяют площадь многоугольных фигур, т. е. выполняется

Теорема II. *Определенная только что площадь обладает свойствами инвариантности и аддитивности:*

1. *Если фигура F имеет определенную площадь $S(F)$, то всякая равная ей фигура F' тоже имеет определенную площадь, и притом равную $S(F)$.*

2. *Если фигура F составлена из фигур F_1, F_2 с определенными площадями $S(F_1)$, $S(F_2)$, то она тоже имеет определенную площадь $S(F)$ и $S(F) = S(F_1) + S(F_2)$.*

Выразим условие, при котором фигура имеет определенную площадь, несколько иначе.

Если многоугольная фигура H содержится в фигуре F, то многоугольная фигура G, содержащая F, получается прибавлением к H некоторой многоугольной фигуры K — «разности» $G - H$. Эта фигура, очевидно, содержит границу фигуры F (рис. 21, 22). Ее площадь равна разности площадей фигур G и H:

$$S(K) = S(G) - S(H).$$

Стало быть, то, что площади фигур G и H могут быть сколь угодно близки и тем самым фигура F имеет определенную площадь, равносильно тому, что площадь фигуры $G - H$ может быть сколь угодно малой. То есть граница данной фигуры F может быть заключена в многоугольную фигуру сколь угодно малой площади. И можно сформулировать: фигура имеет определенную площадь тогда и только тогда, когда ее границу можно заключить в многоугольную фигуру сколь угодно малой площади. Но если фигура может быть заключена в многоугольную фигуру сколь угодно малой площади, то ее площадь равна нулю (как, например, равна нулю площадь отрезка).

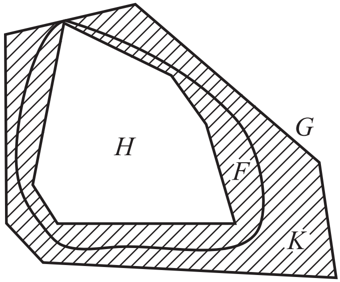


Рис. 22.

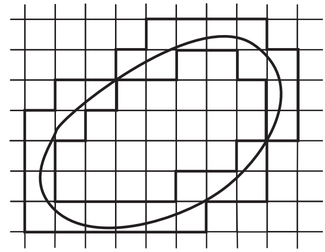


Рис. 23.

Это позволяет выразить полученное условие существования у фигуры определенной площади так.

Теорема 1. *Фигура имеет определенную площадь тогда и только тогда, когда площадь ее границы равна нулю.*

И можно пересказать теорему II;

Теорема II а. *Площадь фигур с определенной площадью обладает теми же свойствами* 1–4, *как, в частности, площадь многоугольных фигур.*

Эта теорема будет доказана в следующих параграфах.

К определению площади можно подойти, исходя из способа ее измерения с помощью квадратных сеток (рис. 23). Это обобщает тот прием, каким находят площадь прямоугольника или измеряют площадь практически с помощью палетки. Коротко можно сказать:

Площадь фигуры можно определить как величину, измеряемую площадью квадратов сетки, содержащихся в фигуре и покрывающих фигуру, если у этих чисел есть общий предел.

Это определение одинаково для многоугольных и других фигур и равносильно данным выше. На его основе и будут дальше доказаны теоремы I, II.

Замечание 1. То условие, что фигура имеет определенную площадь, если площадь ее границы равна нулю, мало наглядно и, что более существенно, *само использует понятие площади.* Однако заменить его другим не удается, и можно указать только те или иные более частные условия, когда оно выполняется. Следующее условие будет все же достаточно общим.

Назовем *криволинейным отрезком* такую фигуру — кривую, — которая в подходящих координатах представляется уравнением $y = f(x)$ с непрерывной функцией f, заданной на каком-либо замкнутом промежутке. В § 8 будет доказана

Теорема III. *Всякая фигура, ограниченная конечным числом криволинейных отрезков, имеет определенную площадь; для таких фигур выполняется все то, что было сказано о площади многоугольных фигур.*

Замечание 2. Приведем пример фигуры без определенной площади. Представим себе прямоугольник, составленный из двух квадратов P, Q с общей стороной. Пусть на одной стороне квадрата P введена координата x. Представим себе также фигуру R, состоящую из всех лежащих в этом квадрате отрезков, параллельных другой его стороне и имеющих концы в точках с рациональными значениями координаты x. Эти отрезки плотно «штрихуют» весь квадрат. Фигура F, составлен-

ная из квадрата Q и фигуры R, содержится в прямоугольнике $P+Q$, содержит квадрат Q, но большей многоугольной фигуры не содержит. Эта фигура F не имеет определенной площади.

В этом примере фигура F имеет часть R довольно необычного «патологического» строения. Но можно привести примеры областей, которые не имеют определенной площади. Эти примеры строятся не так просто, и мы их здесь приводить не будем. Полезно, однако, знать, что даже не всякая область имеет определенную площадь в смысле принятого выше определения.

Это можно понять практически. Представим себе земельный участок, ограниченный с одной стороны оврагом с сильно изрезанным краем. Для владельца участка зигзаги края неудобны, и площадь участка он будет измерять без них. Но землеустроитель может настаивать на том, чтобы учитывать и площадь «зигзагов». Таким образом, практически оказывается, что площадь можно оценивать поразному. Математическая идеализация и приводит к областям, не имеющим определенной площади из-за особенностей границы.

Замечание 3. Подобно тому, как для многоугольных фигур площадь есть функция со свойствами 1–4, можно сказать, что, вообще, площадь других фигур — это функция с теми же свойствами. Однако для того чтобы функция была определена, нужно указать область ее задания, в нашем случае — множество тех фигур, для которых она определена, но характеризуется другими свойствами, о чем мы здесь говорить не будем.

§ 2. Определение площади измерением

Представим себе, что плоскость разбита прямыми на единичные квадраты подобно клетчатой бумаге. Эти квадраты, в свою очередь, разбиты на меньшие равные квадраты, те на еще меньшие и т. д. Так мы представляем себе последовательность все более уменьшающихся квадратных сеток, покрывающих плоскость. Эти сетки мы перенумеруем: первая, состоящая из единичных квадратов, вторая, третья и т. д.

Единичному квадрату и, соответственно, всем квадратам первой сетки припишем площадь, равную единице:

$$S(E_1) = 1.$$

Если в n-й сетке единичный квадрат разделен на N_n квадратов, то

каждому квадрату n-й сетки приписываем площадь

$$S(E_n) = \frac{1}{N_n}.$$

Фигуре, составленной из квадратов сетки, приписывается площадь, равная сумме их площадей. (Мы говорим здесь «площадь» для краткости, хотя имеем в виду численную площадь; строго говоря, пока речь идет о числах, приписываемых фигурам в качестве их численной площади.)

Пусть теперь F — какая-либо фигура. Сопоставим ей две фигуры из квадратов n-й сетки: F_n^i, F_n^e (i, e — первые буквы слов interior — внутренний, exterior — внешний). Фигура F_n^i состоит из всех квадратов n-й сетки, внутренности которых содержатся в F (рис. 23). Фигура F_n^e состоит из всех квадратов n-й сетки, внутри которых есть точки из F (рис. 23). Площади этих фигур обозначаем $S(F_n^i)$, $S(F_n^e)$. Очевидно, каждый квадрат из F_n^i входит в F_n^e, и поэтому

$$F_n^i \subset F_n^e, \quad S(F_n^i) \leqslant S(F_n^e). \tag{1}$$

Перейдем к $(n+1)$-й сетке. Каждый квадрат n-й сетки, входящий в фигуру F_n^i, разобьется на квадраты $(n+1)$-й сетки. Внутренности их будут также содержаться в фигуре F, и, стало быть, все они будут входить в фигуру F_{n+1}^i. Поэтому $F_{n+1}^i \supset F_n^i$ (к квадратам из F_n^i могут прибавиться квадраты, содержащиеся в квадратах из F_n^e). Соответственно

$$S(F_{n+1}^i) \geqslant S(F_n^i). \tag{2}$$

Квадраты, входящие в фигуру F_{n+1}^e, т. е. содержащие внутри точки из F, будут, очевидно, содержаться в квадратах, составляющих фигуру F_n^e. Поэтому $F_{n+1}^e \subset F_n^e$ и, соответственно,

$$S(F_{n+1}^e) \leqslant S(F_n^e). \tag{3}$$

Таким образом, переходя от первой сетки ко второй и т. д., мы получаем две последовательности чисел $S(F_n^i)$ и $S(F_n^e)$. При этом ввиду (2)

$$S(F_1^i) \leqslant S(F_2^i) \leqslant S(F_3^i) \leqslant \ldots$$

Эта последовательность ограничена (так как все квадраты фигур F_n^i содержатся в единичных квадратах, покрывающих фигуру F).

Таким образом, последовательность чисел $S(F_n^i)$ неубывающая и ограниченная. Поэтому она имеет предел; обозначим его $S_1(F)$, т.е. положим

$$S_i(F) = \lim_{n\to\infty} S(F_n^i). \tag{4}$$

Это число $S_i(F)$ мы назовем *внутренней площадью фигуры* F.

Рассмотрим теперь последовательность чисел $S(F_n^e)$. Из (3) следует, что

$$S(F_1^e) \geqslant S(F_2^e) \geqslant S(F_3^e) \geqslant \ldots$$

Таким образом, числа $S(F_n^e)$ образуют невозрастающую последовательность, и притом ограниченную, так как все эти числа $\geqslant 0$. Поэтому последовательность имеет предел; обозначим его $S_e(F)$, т.е. положим

$$S_e(f) = \lim_{n\to\infty} S(F_n^e). \tag{5}$$

Это число $S_e(F)$ назовем *внешней площадью фигуры* F.

Внешняя площадь всякой фигуры не меньше внутренней:

$$S_e(F) \geqslant S_i(F). \tag{6}$$

Действительно, согласно (1) $S(F_n^e) \geqslant S(F_n^i)$ при всяком n. Поэтому такое же неравенство будет между пределами этих чисел, т.е. выполняется (6).

Определение. Если внешняя и внутренняя площади фигуры равны, т.е. $S_e(F) = S^i(F)$, то их общее значение принимается за площадь фигуры F, и мы говорим: фигура F имеет определенную площадь $S(F)$:

$$S(F) = S_e(F) = S_i(F). \tag{7}$$

Таким образом, можно дать определение:

Площадью (*численной площадью*) фигуры называется общее значение ее внешней и внутренней площади, когда они равны.

Вначале фигурам, составленным из квадратов какой-либо сетки, была приписана площадь — сумма площадей этих квадратов. Нужно убедиться, что это согласуется с данным общим определением площади.

Пусть фигура F состоит из квадратов m-й сетки; тем самым $F = F_m^i = F_m^e$ и при всех $n > m$ также $F = F_n^i = F_n^e$. Поэтому при всех $n > m$ площадь, первоначально приписанная фигуре F, будет

$$S(F) = S(F_n^i) = S(F_n^e).$$

В пределе при $n \to \infty$ получим

$$S(F) = S_i(F) = S_e(F),$$

т. е. $S(F)$ является общим значением внутренней и внешней площади. Тем самым $S(F)$ является площадью фигуры F в смысле данного общего определения.

Итак, мы сопоставили некоторым фигурам числа $S(F)$ по формуле (7) и назвали их площадью. Но нужно еще доказать, что эти числа действительно обладают свойствами, характеризующими площадь, и выяснить, для каких фигур эти числа определены. Короче, нужно доказать, что они — те, о которых говорится в теоремах I, II. Это мы дальше и докажем. Попутно мы установим также основные свойства внутренней и внешней площади.

§ 3. Аддитивность площади

Здесь мы докажем теорему об аддитивности площадей:

Теорема 1. *Если фигуры F_1, F_2 имеют определенную площадь и не имеют общих внутренних точек, то фигура $F_1 + F_2$ тоже имеет определенную площадь и*

$$S(F_1 + F_2) = S(F_1) + S(F_2). \tag{1}$$

Эта теорема будет получена как следствие теоремы о «полуаддитивности» внешней и внутренней площади.

Теорема 2. *Для любых двух фигур F_1, F_2 без общих внутренних точек*

$$\begin{gathered} S_e(F_1 + F_2) \leqslant S_e(F_1) + S_e(F_2), \\ S_i(F_1 + F_2) \geqslant S_i(F_1) + S_i(F_2). \end{gathered} \tag{2}$$

(Такие же соотношения, как (1), (2), выполняются для любого числа слагаемых F_1, F_2, ..., F_k, что выводится из (1), (2) обычным путем.)

Доказательство теоремы 2. Для внешней площади выполняется общее утверждение:

Лемма 1. *Для любых фигур F_1, F_2 и их объединения $F_1 \cup F_2$*

$$S_e(F_1 \cup F_2) \leqslant S_e(F_1) + S_e(F_2).$$

Доказательство. Пусть $F = F_1 \cup F_2$. Если квадрат n-й сетки входит в F_n^e, то это значит, по условию, что он содержит внутри себя

точки из F. Тем самым он содержит внутри себя точки хотя бы одной из фигур F_1, F_2, скажем, фигуры F_1. Это значит, что он входит в фигуру $F^e_{1,n}$.

Итак, всякий квадрат фигуры F^e_n принадлежит хотя бы одной из фигур $F^e_{1,n}$, $F^e_{2,n}$. Площади же всех этих фигур слагаются из площадей составляющих их квадратов. Тем самым

$$S(F^e_n) \leqslant S(F^e_{1,n}) + S(F^e_{2,n}).$$

Переходя к пределу при $n \to \infty$, получим

$$S_e(F) \leqslant S_e(F_1) + S_e(F_2). \quad \square$$

Лемма 2. *Если фигуры F_1, F_2 не имеют общих внутренних точек, то*

$$S_i(F_1 + F_2) \geqslant S_i(F_1) + S_i(F_2).$$

Доказательство. Пусть фигура F составлена из F_1 и F_2:

$$F = F_1 + F_2.$$

Пусть какой-то квадрат Q n-й сетки содержится в фигуре $F^i_{1,n}$, т. е. внутренность его содержится в F_1. Тем самым она содержится в F, и, значит, Q содержится в фигуре F^i_n. Вместе с тем внутренность этого квадрата не содержится внутри F_2, так как F_2 не имеет с F_1 общих внутренних точек. Тем самым квадрат Q не содержится в фигуре $F^i_{2,n}$. Аналогично, если квадрат содержится в $F_{2,n}$, то он не содержится в $F_{1,n}$. Таким образом, мы видим, что фигуры $F^i_{1,n}$, $F^i_{2,n}$ не имеют общих квадратов и все их квадраты входят в фигуру F_n. Тем самым для сумм их площадей получается, что

$$S(F^i_{1,n}) + S(F^i_{2,n}) \leqslant S(F^i_n).$$

Переходя к пределу при $n \to \infty$, получаем

$$S_i(F_1) + S_i(F_2) \leqslant S_i(F) = S_i(F_1 + F_2).$$

Таким образом, лемма 2 доказана. $\square$

Вместе с леммой 1 они дают теорему 2. $\square$

Доказательство теоремы 1. Теорему 1 можно сформулировать так:

Если фигура F составлена из фигур F_1, F_2 с определенной площадью, то она тоже имеет определенную площадь и

$$S(F) = S(F_1) + S(F_2).$$

Доказательство. Пусть фигура F составлена из фигур F_1, F_2 с определенной площадью, т. е. $S_e(F_1) = S_i(F_1) = S(F_1)$, и то же для F_2. Тогда из лемм 1, 2

$$S_e(F) \leqslant S(F_1) + S(F_2),$$
$$S_i(F) \geqslant S(F_1) + S(F_2).$$

А так как заведомо $S_i(F) \leqslant S_e(F)$, то получается, во-первых, что

$$S_e(F) = S_i(F),$$

т. е. фигура F имеет определенную площадь $S = S_i = S_e$. И, во-вторых,

$$S(F) = S(F_1) + S(F_2).$$

Теорема 1 доказана. $\square$

Монотонность площади.

Теорема 3. *Если $F \subset F'$, то*

$$S_i(F) \geqslant S_i(F'), \quad S_e(F) \geqslant S_e(F').$$

Поэтому если фигуры F, F' имеют определенные площади (т. е. у них $S = S_e = S_i$)), то

$$S(F) \geqslant S(F').$$

Доказательство. Пусть $F' \supset F'$. Тогда в n-й сетке всякий квадрат, содержащий внутри себя точки из F', тем самым содержит внутри себя точки из F. Следовательно, $F_n'^e \subset F_n^e$ и

$$S(F_n'^e) \leqslant S(F_n^e).$$

Переходя к пределу (при $n \to \infty$), получим:

$$S_e(F') \leqslant S_e(F).$$

Вывод для S_i аналогичен. Всякий открытый квадрат, содержащийся в F', содержится и в F. Поэтому $F_n'^i \subset F_n^i$, так что $S(F_n'^i) \leqslant S(F_n')$ и в пределе $S_i(F') \leqslant S(F)$, что и требовалось доказать. $\square$

§ 4. Фигуры с определенной площадью

Обозначение. Здесь для обозначения границы воспользуемся символом ∂ (а не fr).

Теорема 1. *Всякая многоугольная фигура имеет определенную площадь.*

Эта теорема вытекает из следующей общей теоремы.

Теорема 2. *Фигура имеет определенную площадь тогда и только тогда, когда внешняя площадь ее границы равна нулю.*

Так как всегда $S_e \geqslant S_i \geqslant$, то из $S_e(F) = 0$ следует $S_i(F) = 0$, т. е. $S_e(F) = S_i(F)$, и, значит, фигура F имеет определенную площадь, равную нулю. Поэтому теорему 2 можно формулировать и так:

Теорема 2а. *Фигура имеет определенную площадь тогда и только тогда, когда площадь ее границы равна нулю.*

Доказательство теоремы 2. Теорема 2 вытекает из следующего изящного утверждения.

Лемма 1. *Для всякой фигуры F*

$$S_e(F) - S_i(F) = S_e(\partial F),$$

где ∂F — граница F.

Доказательство. Пусть F — данная фигура и ∂F — ее граница. Рассмотрим в данной n-й сетке фигуры F_n^e, F_n^i, $(\partial F)_n^e$. Покажем, что фигура $(\partial F)_n^e$ состоит из тех квадратов фигуры F_n^e, которые не входят в F_n^i (рис. 23).

В самом деле, если квадрат Q входит в F_n^e, но не в F_n^i, то это значит, что внутри него есть точки из F, но также точки, не принадлежащие F (иначе он содержался бы в F_n^i). Следовательно, внутри этого квадрата есть точки границы F. Тем самым такой квадрат входит в $(\partial F)_n^e$.

Пусть теперь Q — квадрат, входящий в $(\partial F)_n^e$. Он содержит внутри себя точки из ∂F, т. е. точки границы фигуры F. Тем самым он содержит внутри себя как точки, принадлежащие F, так и точки, не принадлежащие F. А это значит, что он входит в фигуру F_n^e, но не в F_n^i.

Таким образом, фигура $(\partial F)_n^e$ образуется всеми квадратами, которые входят в F_n^e, но не в F_n^i. Поэтому для площадей выполняется равенство:

$$S(F_n^e) - S(F_n^i) = S((\partial F)_n^e).$$

Переходя к пределу, получаем

$$S_e(F) - S_i(F) = S_e(\partial F),$$

что и требовалось доказать. □

Из доказанной леммы сразу следует теорема 2.

Доказательство. То, что фигура F имеет определенную площадь, означает по определению, что $S_e(F) = S_i(F)$. А из доказанной леммы следует, что это равносильно

$$S_e(\partial F) = 0.$$

Теорема 2 доказана. □

Доказательство теоремы 1. Теорема 1 о том, что многоугольная фигура имеет определенную площадь, выводится из доказанной теоремы 2. Для этого докажем три леммы.

Лемма 2. *Если a — длина стороны квадрата данной сетки, то отрезок длины l пересекает не более $2l/a + 4$ квадратов.*

Доказательство. Пусть дана квадратная сетка с длиной стороны квадрата, равной a. Сетка образована прямыми двух взаимно перпендикулярных направлений. Любой данный отрезок образует с ними углы, в сумме равные $90°$, и, стало быть, с прямыми одного из этих направлений — угол, не больший $45°$. Такие прямые назовем *горизонтальными*, другие — *вертикальными*.

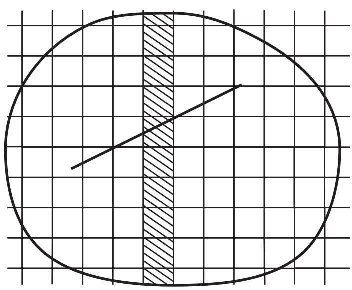


Рис. 24

Вертикальные прямые делят плоскость на полосы ширины a. Отрезок, пересекая такую полосу, может пересекать, самое большее, два квадрата, иначе его угол с горизонтальной прямой был бы больше $45°$ (рис. 24).

Если длина отрезка равна l, то он пересекает целиком никак не больше l/a полос. Вместе с ними он пересекает, самое большее, $2l/a$ квадратов.

У каждого из концов он может пересекать по одной полосе и, значит, еще по 2 квадрата. Так что в целом он пересекает не более $2l/a+4$ квадратов, что и требовалось доказать. □

Лемма 3. *Площадь отрезка равна нулю.*

Доказательство. Пусть фигура F представляет собой отрезок длины l. Пусть a_n — длина стороны квадрата n-й сетки. Тогда, по доказанной лемме 2, отрезок пересекает не более $2l/a_n+4$ квадратов сетки. Значит, фигура F_n^e состоит не более чем из такого числа квадратов. Поскольку площадь одного квадрата равна a_n^2, то площадь фигуры F_n^e оценивается так:

$$S(F_n^e) \leqslant \left(2\frac{l}{a_n}+4\right)a_n^2 = 2la_n + 4a_n^2.$$

Когда $n \to \infty$, то $a_n \to 0$, а $S(F_n^e) \to S_e(F)$. Следовательно, $S_e(F)=0$, что и требовалось доказать. □

Лемма 4. *Площадь объединения конечного числа фигур нулевой площади равна нулю.*

Доказательство. Если фигура F служит объединением фигур $F_1, F_2, \ldots, F_m$, то по лемме 1 § 3

$$S_e(F) \leqslant S_e(F_1)+S_e(F_2)+\ldots+S_e(F_m).$$

Поэтому если все фигуры $F_1, \ldots, F_m$ нулевой площади, то и $S_e(F)=0$. □

Доказательство теоремы 1 получается теперь сразу. Граница многоугольной фигуры состоит из конечного числа отрезков, а потому, как следует из лемм 4, 3, площадь ее равна нулю. Следовательно, согласно теореме 2, многоугольная фигура имеет определенную площадь. □

§ 5. Площади равных многоугольных фигур

Теорема 1. *У равных многоугольных фигур площади равны.*

Доказательство проходит в три этапа.

Равенство площадей для параллельно перенесенных фигур.

Лемма 1. *Площадь прямоугольника со сторонами, параллельными линиям квадратной сетки, равна произведению длин его сторон.*

Доказательство получаем обычным путем, как в школьном курсе, беря квадраты, содержащиеся в данном прямоугольнике F и покрывающие его, т. е. беря фигуры F_n^i и F_n^e. (Правда, в школьном курсе квадраты «накладывают» от двух сторон прямоугольника, тут же они, вообще говоря, располагаются иначе. Но это несущественно.) □

Лемма 2. *Фигуры, составленные из равных прямоугольников со сторонами, параллельными линиям сетки, имеют равные площади.*

Доказательство. Из леммы 1 следует, что равные прямоугольники, о которых тут говорится, имеют равные площади. Следовательно, по аддитивности площади, и фигуры, составленные из них, имеют равные площади, что и требовалось доказать. □

Лемма 3. *Для многоугольных фигур F всегда $F_n^i \subset F \subset F_n^e$.*

Это очевидно. □

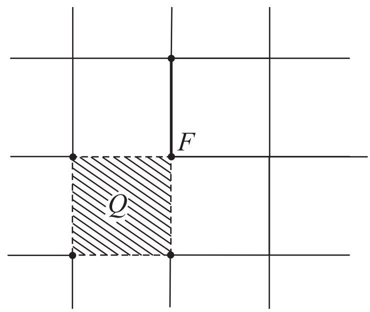


Рис. 25

Замечание. Совершенно так же имеет место общее утверждение: если фигура F состоит из внутренности и ее границы (т. е. фигуру F ограничивает ее граница ∂F; см. ч. II, гл. 2, § 2), то $F_n^i \subset F \subset F_n^e$. Но если фигура имеет другое строение, то эти включения могут не выполняться. Простейший пример: если F — внутренность квадрата Q n-й сетки с добавленной стороной другого квадрата, то $F_n^i = F_n^e = Q$, но $Q \not\subset F$ и $F \not\subset Q$ (рис. 25).

Лемма 4. *Если одна многоугольная фигура получается из другой параллельным переносом, то их площади равны.*

Доказательство. Пусть F, F' — многоугольные фигуры, и пусть $F' = tF$, где t — параллельный перенос. Возьмем в n-й сетке фигуры F_n^i, F_n^e. По лемме 3 имеем $F_n^i \subset F \subset F_n^e$. Поэтому также $tF_n^i \subset tF \subset tF_n^e$. Отсюда по монотонности площади

$$S(tF_n^i) \leqslant S(tF) \leqslant S(tF_n^e).$$

Фигуры F_n^i, F_n^e состоят из прямоугольников со сторонами на прямых сетки, поэтому согласно лемме 2

$$S(tF_n^i) = S(F_n^i), \quad S(tF_n^e) = S(F_n^e).$$

Благодаря этому предыдущие неравенства можно переписать так:

$$S(F_n^i) \leqslant S(tF) \leqslant S(F_n^e).$$

При $n \to \infty$ здесь крайние члены сходятся к $S(F)$, стало быть,

$$S(tF) = S(F),$$

что и требовалось доказать. □

Квадраты каждой из наших сеток имеют одну и ту же определенную площадь и получаются друг из друга параллельными переносами. Если подвергнуть их какому-либо перемещению, то получатся опять сетки из равных и параллельно расположенных квадратов. Поэтому из леммы 4 вытекает:

Следствие. *Квадратные сетки, получаемые из данных путем какого угодно перемещения, состоят каждая из квадратов одной и той же площади. Площадь каждого квадрата составляет поэтому такую же долю основного квадрата, как в данных сетках.* □

(Но пока еще неизвестно, равны ли их площади площадям квадратов исходной сетки. Это надо доказать.)

Равенство площадей многоугольных фигур, получаемых отражением.

Лемма 5. *Если фигура F' получается из многоугольной фигуры F отражением в прямой, то она имеет ту же площадь, что и F.*

Доказательство. Пусть F — данная многоугольная фигура. Произведем отражение в какой-нибудь прямой; обозначая его r, имеем фигуру $F' = rF$. Вместе с фигурой F подвергнем тому же отражению все наши квадратные сетки. Получим, как указано в следствии леммы 4, такие же сетки, только неизвестно, будет ли основной квадрат единичным. Если его площадь равна k, то все квадраты отраженных сеток будут тоже отличаться в k раз от площади соответствующих квадратов исходных сеток.

Фигуры F_n^i, F_n^e перейдут при этом в фигуры rF_n^i, rF_n^e, построенные по n-й отраженной сетке так же, как F_n^i, F_n^i построены по исходной сетке. Поэтому отношения их площадей к площадям основных квадратов будут одни и те же, так что

$$S(rF_n^i) = kS(F_n^i), \quad S(rF_n^e) = kS(F_n^e).$$

Так как, по лемме 3, $F_n^i \subset F \subset F_n^e$, то точно так же $rF_n^i \subset rF \subset rF_n^e$. Поэтому для площадей получаем:

$$S(rF_n^i) \leqslant S(rF) \leqslant S(rF_n^e),$$

или, принимая во внимание предыдущие равенства:

$$kS(F_n^i) \leqslant S(rF) \leqslant kS(F_n^e).$$

При переходе к пределу при $n \to \infty$ площади $S(F_n^i)$, $S(F_n^e)$ сходятся к $S(F)$, и потому получаем:

$$kS(F) \leqslant S(rF) \leqslant kS(F).$$

Следовательно,

$$S(rF) = kS(F). \tag{1}$$

Весь этот вывод применим к любой многоугольной фигуре и, значит, к rF. Поэтому, подставляя в последнее равенство rF вместо F, получим

$$S(rrF) = kS(rF) = k^2 S(F).$$

Но повторное отражение дает тождественное преобразование, так что $rrF = F$. Поэтому $k^2 = 1$, и из (1) получаем, что

$$S(rF) = S(F),$$

что и требовалось доказать. □

Доказательство теоремы 1. Теорема 1 о равенстве площадей равных многоугольных фигур непосредственно следует из доказанной леммы 5. Действительно, фигура, равная данной, получается из нее некоторым перемещением, а всякое перемещение может быть получено как результат нескольких (не более 3) последовательных отражений (теорема 1.3.5, ч. 3). При каждом отражении многоугольная фигура, согласно лемме 5, переходит в многоугольную фигуру с той же площадью. Поэтому то же будет и при нескольких отражениях. Тем самым фигура, равная данной, имеет ту же площадь, что и требовалось доказать. □

§ 6. Окончание доказательства теоремы I

В § 1 была высказана теорема I:

Пусть выбран некоторый квадрат E. Тогда каждой многоугольной фигуре F может быть отнесена, и притом единственная, численная площадь, т. е. такое число $S(F)$, что будут выполнены следующие четыре условия:

1) $S(F) > 0$,
2) *если* $F \cong F'$, *то* $S(F) = S(F')$ (*инвариантность*),
3) $S(F_1 + F_2) = S(F_1) + S(F_2)$ (*аддитивность*),
4) $S(E) = 1$.

Если вместо квадрата E взять другой E', то соответствующие числа S' отличаются от S только общим множителем:

$$S'(F) = kS(F); \quad k = S'(E) = \frac{1}{S(E')}. \tag{1}$$

В § 2–5 мы определили численные площади — числа $S(F)$, — в частности для многоугольных фигур, и доказали, что они действительно обладают свойствами 2, 3. Тем самым доказано существование численной площади у многоугольных фигур.

Остается доказать ее единственность при данном квадрате E и правило ее пересчета (1) для другого квадрата E'.

Единственность численной площади.

Теорема 1. *При данном единичном квадрате E численная площадь многоугольных фигур определяется однозначно. То есть если многоугольным фигурам сопоставлены числа S со свойствами* 1–4 *теоремы* I, *то они представляют собой численные площади, определенные с помощью квадратных сеток.*

Доказательство. Допустим, многоугольным фигурам отнесены числа $S'(F)$ со свойствами 1–4. Нужно доказать, что $S'(F) = S(F)$. По свойству 4, $S'(E) = 1$, и из инвариантности следует, что в каждой нашей сетке квадратам отвечает одно и то же значение $S'(E_n)$. Если в единичном квадрате их N_n штук, то $N_n S'(E_n) = S'(E) = 1$. Но точно так же, по нашему определению численной площади, $N_n S(E_n) = 1$. Стало быть,

$$S'(E_n) = S(E_n).$$

Поэтому при любой фигуре F

$$S'(F_n^i) = S(F_n^i), \quad S'(F_n^e) = S(F_n^e). \tag{2}$$

Для многоугольной фигуры F

$$F_n^i \subset F \subset F_n^e,$$

и «разности» $F - F_n^i$ и $F_n^e - F$ представляют собой многоугольные фигуры (или пустое множество, если $F = F_n^i$); по аддитивности

$$S'(F) = S'(F_n^i) + S'(F - F_n^i),$$

а по свойству 1 имеем $S'(F - F_n^i) \geqslant 0$. Поэтому $S'(F) \geqslant S'(F_n^i)$. Аналогично заключаем, что $S'(F) \leqslant S'(F_n^e)$. Таким образом, в силу равенств (2) получаем, что

$$S(F_n^i) \leqslant S'(F) \leqslant S(F_n^e).$$

При $n \to \infty$ крайние члены сходятся к $S(F)$, и, стало быть, $S'(F) = S(F)$, что и требовалось доказать. $\square$

Замена единицы площади.

Теорема 2. *При замене единицы площади все численные ее значения для многоугольных фигур умножаются на один и тот же множитель. Если E' — «новый» единичный квадрат и S' — определенная по нему численная площадь (т. е. когда $S'(E') = 1$), то для фигуры F*

$$S'(F) = kS(F), \quad k = S'(E) = \frac{1}{S(E')}.$$

Доказательство. Пусть E' — «новый» единичный квадрат и $k = S'(E)$ — численная площадь», какую получает при этом «первоначальный» единичный квадрат E. «Новая» численная площадь всякой фигуры G, состоящей из квадратов n-й сетки, построенной по E, будет равна сумме площадей S' составляющих ее квадратов (согласно доказанным свойствам численной площади). Если квадрат E_n n-й сетки составляет $1/N_n$ квадрата E, а число квадратов в фигуре G равно M_n, то

$$S'(G) = M_n \frac{S'(E)}{N_n} = S(G)S'(E),$$

так как $M_n/N_n = S(G)$ (по определению).

Поэтому для фигур F_n^i, F_n^e, построенных для любой фигуры F,

$$S'(F_n^i) = S(F_n^i)S'(E),$$
$$S'(F_n^e) = S(F_n^e)S'(E).$$

Для многоугольной фигуры (по лемме 3, § 5) $F_n^i \subset F \subset F_n^e$. Поэтому

$$S'(F_n^i) \leqslant S'(F) \leqslant S'(F_n^e),$$

и из предыдущих равенств

$$S(F_n^i)S'(E) \leqslant S'(F) \leqslant S(F_n^e)S'(E).$$

При $n \to \infty$ величины $S(F_n^i)$, $S(F_n^e)$ сходятся к $S(F)$. Поэтому в пределе получаем

$$S(F)S'(E) \geqslant S'(F) \geqslant S(F)S'(E),$$

т. е. $S'(F) = S(F)S'(E)$, что и требовалось доказать. $\square$

С доказательством теорем 1, 2 теорема I доказана полностью. $\square$

§ 7. Площадь немногоугольных фигур: теоремы II, IIa

В § 1 были высказаны две теоремы о площади любых фигур, у которых площадь границы равна нулю. Теорема IIa говорит, что численная площадь таких фигур удовлетворяет тем же условиям, что и площадь многоугольных фигур. Теорема II утверждает, что площади этих фигур можно находить, как находят площадь круга, по площадям «вписанных» и «описанных» многоугольных фигур. Здесь мы докажем сначала теорему IIa, а потом — теорему II, и установим еще одно важное свойство фигур с определенной площадью.

Нахождение площади посредством многоугольных фигур. Установим сначала общее условие существования площади.

Теорема 1a. *Если для фигуры F существуют такие фигуры $G \supset F$ и $H \subset F$ с определенными площадями, что разности их площадей сколь угодно малы, то F имеет определенную площадь и она равна общему пределу площадей $S(G)$, $S(H)$.*

Доказательство. Если $G \supset F \supset H$ и фигуры G, H имеют площади $S(G)$, $S(H)$, то (по теореме 3, § 3)

$$S(G) \geqslant S_e(F) \geqslant S_i(F) \geqslant S(H).$$

Поэтому если разность $S(G) - S(H)$ может быть сколь угодно малой, то $S_e(F) = S_i(F)$, т. е. F имеет определенную площадь $S = S_e = S_i$. И беря фигуры G_n, H_n так, что $S(G_n) - S(H_n) \to 0$, получим, что

$$S(G_n) \to S(F), \quad S(H_n) \to S(F).$$

Теорема 1 а доказана. □

В частности, фигуры G, H могут быть многоугольными. Поэтому в доказанной теореме 1a содержится утверждение: *если для фигуры F существуют многоугольные фигуры $G \supset F$ и $H \subset F$ со сколь угодно малой разностью площадей, то F имеет определенную площадь.*

Выполняется также обратное утверждение:

Теорема 1б. *Если фигура F имеет определенную площадь $S(F)$, то существуют такие многоугольные фигуры G_n, H_n ($n = 1, 2, 3, \ldots$), что $G_n \supset F \supset H_n$ и*

$$S(G_n) \to S(F), \quad S(H_n) \to S(F).$$

Доказательство. По принятому здесь определению площади

$$S(F) = \lim_{n\to\infty} S(F_n^i) = \lim_{n\to\infty} S(F_n^e).$$

Фигура F_n^i состоит из квадратов, внутренности которых содержатся в F. Построив внутри каждого такого квадрата квадрат, к нему достаточно близкий, можно получить многоугольную фигуру $H_n \subset F$ с площадью, как угодно близкой к $S(F_n^i)$.

Получим такую последовательность многоугольных фигур $H_n \subset F$, что $S(H_n) \to S(F)$.

Фигура F_n^e состоит из квадратов, внутри которых содержатся точки из F. Но у фигуры F могут быть еще точки, лежащие на сторонах квадратов сетки. Все такие стороны можно окружить прямоугольниками сколь угодно малой площади. Присоединив их к фигуре F_n^e, получим многоугольную фигуру G_n с площадью, сколь угодно близкой к F_n^e. Получим такую последовательность многоугольных фигур $G_n \subset F$, что $S(G_n) \to S(F)$. Теорема 1б доказана. $\square$

Теорема 1б вместе о теоремой 1а дают:

Теорема 1. *Фигура F имеет определенную площадь в смысле определения с помощью квадратных сеток тогда и только тогда, когда существуют как содержащие ее, так и содержащиеся в ней многоугольные фигуры G и H ($G \supset F \supset H$) со сколь угодно малой разностью площадей $S(G) - S(H)$. Площадь фигуры F равна при этом общему пределу площадей $S(G)$, $S(H)$.* $\square$

Равенство площадей равных фигур.

Теорема 2. *Если фигура F имеет определенную площадь, то всякая равная ей фигура тоже имеет определенную площадь, и притом ту же самую.*

Доказательство. Пусть F — фигура с определенной площадью $S(F)$, а F' — фигура, ей равная, т. е. полученная из F некоторым перемещением d: $F' = dF$. По теореме 1б существуют такие G_n, H_n — многоугольные фигуры: $G_n \supset F \supset H_n$, что

$$S(G_n) \to S(F), \quad S(H_n) \to S(F). \tag{1}$$

Так как $G_n \supset F \supset H_n$, то для фигур G'_n, H'_n, полученных из G_n, H_n тем же перемещением d, будет

$$G'_n \supset F' \supset H'_n.$$

Поэтому

$$S(G'_n) \geqslant S_e(F') \geqslant S_i(F') \geqslant S(H'_n). \tag{2}$$

Но так как фигуры G'_n, H'_n равны многоугольным фигурам G_n, H_n, то по теореме 1, § 5 их площади те же. Поэтому предыдущие неравенства

(2) можно переписать так:

$$S(G_n) \geqslant S_e(F') \geqslant S_i(F') \geqslant S(H_n).$$

По соотношениям (1) крайние члены здесь, при $n \to \infty$, стремятся к одному и тому же пределу $S(F)$. Поэтому в пределе получаем

$$S(F) \geqslant S_e(F') \geqslant S_i(F') \geqslant S(F),$$

т. е. $S_e(F') = S_i(F') = S(F)$. Это значит, что фигура F' имеет определенную площадь и притом равную $S(F)$, что и требовалось доказать. □

Совокупность фигур определенной площади. Следующая теорема дает основание получать из фигур определенной площади другие такие фигуры.

Теорема 3. *Объединение и пересечение любого конечного числа фигур определенной площади, как и разность двух таких фигур, всегда оказывается фигурой с определенной площадью* (здесь «разность» фигур F, F' — в обозначениях $F \setminus F'$ — это множество всех точек, принадлежащих F, но не принадлежащих F').

Доказательство. Достаточно доказать сказанное для двух фигур: если фигуры F_1, F_2 имеют определенную площадь, то ее имеют также фигуры $F_1 \cup F_2$, $F_1 \cap F_2$, $F_1 \setminus F_2$. Наличие определенной площади равносильно тому, что площадь границы равна нулю. Поэтому теорема сводится к тому, что если у F_1, F_2 площадь границы равна нулю, то то же будет и у фигур $F_1 \cup F_2$, $F_1 \cap F_2$, $F_1 \setminus F_2$.

Нетрудно доказать, что границы трех последних фигур всегда содержатся в объединении границ самих F_1 и F_2. (Это простое доказательство читатель может провести сам.) А по лемме 4, § 4 объединение фигур нулевой площади тоже имеет нулевую площадь.

Таким образом, если $S(\partial F_1) = S(\partial F_2) = 0$, то $S(\partial F_1 \cup \partial F_2) = 0$, а вместе с этим равны нулю площади границ объединения, пересечения и разности. Таким образом, теорема 3 доказана. □

Единственность площади.

Теорема 4. *Площадь фигур, у которых площадь границы равна нулю, определяется однозначно. То есть если для этих фигур заданы числа S' с теми же условиями, как в теореме* I, *лишь с заменой условия 1 на $S' \geqslant 0$, то эти числа совпадают с численной площадью S.*

Доказательство. Для площади многоугольных фигур единственность установлена теоремой 1, § 6, а согласно теореме 1 площади других фигур определяются по площадям многоугольных фигур,

поэтому их площади тоже определяются однозначно. Подробнее это соображение представляется следующим образом.

Пусть фигурам с определенной площадью сопоставлены числа S' с указанными в теореме свойствами. Пусть $F \supset H$, где F, H — фигуры с определенной площадью, тогда, по теореме 3, $F \setminus H$ — тоже фигура с определенной площадью. Стало быть, по аддитивности чисел S',

$$S'(F) = S'(H) + S'(F \setminus H),$$

и так как $S' \geqslant 0$, то $S'(F) \geqslant S'(H)$.

Следовательно, если F — данная фигура, а G, H — такие многоугольные фигуры, что $G \supset F \supset H$, то

$$S'(G) \geqslant S'(F) \geqslant S'(H).$$

Для многоугольных фигур единственность площади была доказана в теореме 1, § 6. Поэтому $S'(G) = S(G)$, $S'(H) = S(H)$, и последние неравенства можно переписать так:

$$S(G) \geqslant S'(F) \geqslant S(H).$$

По теореме 1 фигуры G, H можно взять так, чтобы $S(G)$ и $S(H)$ сколь угодно мало отличались от $S(F)$. Следовательно, $S'(F) = S(F)$. Теорема 4 доказана. □

Итак, мы доказали теорему IIa:

Для фигур с границей нулевой площади однозначно определяется численная площадь со свойствами:

1) $S(F) \geqslant 0$,

2) *если* $F' \cong F$, *то* $S(F') = S(F)$,

3) $S(F_1 + F_2) = S(F_1) + S(F_2)$,

4) $S(E) = 1$.

В § 2 мы определили численную площадь S, причем свойства 1), 4) выполняются для нее непосредственно; свойство 2) установлено теоремой 2, а 3) — теоремой 1 § 3; единственность (однозначность определения) доказана последней теоремой 4.

Таким образом, теорема II полностью доказана. □

Дополнение. Аналогично теореме 1б можно доказать, что:

Для любой фигуры F и многоугольных фигур G, H, где $G \supset F \supset H$, будет

$$S_e(F) = \inf S(G), \quad S_i(F) = \sup S(H).$$

Отсюда аналогично теореме 2 доказывается, что:

Если $F' \cong F$, то $S_e(F') = S_e(F)$, $S_i(F') = S_i(F)$. Сами теоремы 1 и 2 являются частными случаями этих общих утверждений.

§ 8. Еще о фигурах с определенной площадью

В нашем распоряжении есть два критерия того, что фигура имеет определенную площадь. Первый — это то, что площадь границы фигуры равна нулю. Второй состоит в том, что существуют две многоугольные фигуры со сколь угодно малой разностью площадей — содержащая данную фигуру и содержащаяся в ней. Но эти критерии недостаточно наглядны: они не позволяют по виду самой фигуры сразу определить, что она имеет определенную площадь. Поэтому желательно найти пусть и более частные, но наглядные условия; пока мы установили только одно такое условие — что фигура многоугольная. Но его можно значительно обобщить.

Теорема 1. *Фигура имеет определенную площадь, если ее граница является объединением конечного числа фигур (линий), каждая из которых представляется в подходящих координатах уравнением $y = f(x)$ с непрерывной функцией f на каком-нибудь промежутке $[a, b]$* (так называемых криволинейных отрезков).

Отложив пока доказательство этой теоремы, укажем другую.

Теорема 2. *Фигура имеет определенную площадь, если ее граница не имеет внутренних точек*[35] *и существуют две такие пересекающиеся прямые, что всякая прямая, параллельная одной из них, пересекает границу по конечному числу (или даже по счетному множеству) отдельных точек и отрезков (либо вовсе ее не пересекает).*

Если фигура содержит свою границу, то граница заведомо не имеет внутренних точек, так что в этом случае остается только второе условие теоремы.

Условия теоремы выполняются, например, для любых выпуклых фигур и их объединений в конечном числе (многоугольные фигуры сюда включаются).

Однако эту теорему мы не будем здесь доказывать, так как ее доказательство сложно.

Доказательство теоремы 1. Для доказательства теоремы 1 достаточно доказать, что при ее условиях площадь границы фигуры равна нулю. То есть достаточно доказать, что если фигура слу-

[35] Например, граница фигуры, состоящей из параллельных отрезков, плотно расположенных на квадрате, или фигуры, составленной из всех точек квадрата с рациональными координатами, представляет собой квадрат.

жит объединением конечного числа кривых $y = f(x)$ с непрерывными функциями f, то ее площадь равна нулю. А для этого достаточно доказать то же для одной такой кривой, т. е. доказать следующее.

Лемма 1. *Фигура (кривая), представимая в прямоугольных координатах уравнением $y = f(x)$ с непрерывной функцией f, имеет нулевую площадь.*

Доказательство. Пусть кривая F задана уравнением $y = f(x)$ с непрерывной f на некотором промежутке $[a,\ b]$. Функция, непрерывная на замкнутом промежутке, равномерно непрерывна. Это значит, что при любом положительном ε найдется такое $\delta > 0$, что на всяком промежутке длины, не большей δ, колебание функции оказывается меньше ε. Геометрически это означает, что кривая F над отрезком длины не больше δ может быть заключена в прямоугольник с высотой ε; основание его параллельно оси x и не больше δ (рис. 26, a).

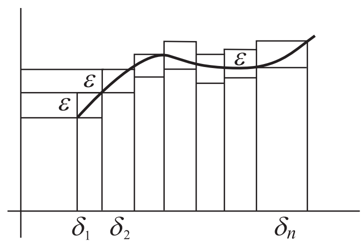


Рис. 26, a

Разделим весь промежуток $[a,\ b]$ на промежутки с длинами $\delta_1, \delta_2, \ldots, \delta_n$, меньшими δ. Тогда кривую F можно включить в фигуру G, состоящую из прямоугольников высотой ε, расположенных над взятыми промежутками. Площадь такого прямоугольника $S_k = \varepsilon\delta_k$. Поэтому площадь всей фигуры G будет

$$S(G) = \sum S_k = \varepsilon \sum \delta_k = \varepsilon(b-a).$$

Так как $F \subset G$, то $S_e(F) \leqslant S(G)$, и, стало быть,

$$S_e(F) \leqslant \varepsilon(b-a).$$

А так как ε можно взять сколь угодно малым, то $S_e(F) = 0$. Таким образом, лемма 1 доказана, а вместе с этим доказана и теорема 1. □

Площадь и интеграл. Между интегралом и площадью есть хорошо известная связь.

Пусть $f(x)$ — непрерывная функция на промежутке $[a,\ b]$. В прямоугольных координатах она представляется кривой $y = f(x)$. Эта кривая F вместе с отрезком $[a,\ b]$ и отрезками, проведенными из его концов параллельно оси y, ограничивает некоторую фигуру T (понятно, что если конец кривой F совпадает с концом отрезка $[a,\ b]$, то

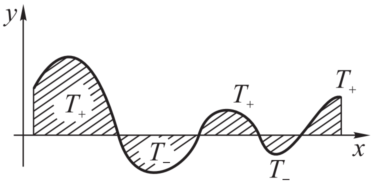


Рис. 26, *б*

отрезок из этого конца не проводится). Если кривая не имеет общих точек с осью x, то эта фигура представляет «криволинейную трапецию». Вообще говоря, фигура T может состоять из двух фигур T_+ и T_- — расположенных над осью x и под ней (рис. 26, *б*). Отрезок $[a,\ b]$ можно отнести и к той, и к другой фигуре. (Тогда фигура T_+ ограничена кривой $y = f(x)$, где $f(x) > 0$, и $y = 0$ там, где $f(x) \leqslant 0$; тут она лежит на оси x. Для T_- — аналогично.)

Согласно теореме 1, фигуры T_+, T_- имеют определенную площадь $S(T_+)$, $S(T_-)$, и вместе с тем

$$\int_a^b f(x)dx = S(T_+) - S(T_-).$$

Доказательство существования интеграла у непрерывной функции совпадает с доказательством существования площади у фигур такого вида, как T_+ и T_-. Говоря геометрически, рассматриваются ступенчатые фигуры, заключающие такую фигуру T и заключенные в ней (рис. 27), и доказывается, что их площади имеют общий предел, когда промежутки, на которые разбивается промежуток $[a,\ b]$, безгранично измельчаются. Разность таких фигур образует многоугольную фигуру, заключающую ту кривую, которая вместе с осью x ограничивает T, и доказательство сводится к тому, что площади таких фигур стремятся к нулю, т. е. площадь кривой F равна нулю. Именно это и было доказано выше в лемме 1.

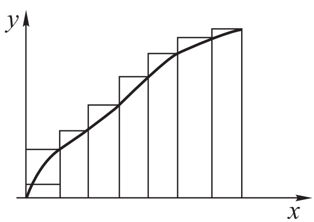


Рис. 27

Таким образом, существование интеграла непрерывной функции и существование площади у соответствующей фигуры T — это одно и то же, и, доказав выше существование ее площади, мы фактически доказали существование интеграла непрерывной функции.

§ 9. Объем

Здесь мы будем рассматривать ограниченные фигуры в пространстве любого числа измерений. Будем говорить, как о фигурах на плоскости, что фигура составлена из нескольких фигур, если она служит их объединением и никакие две из них не имеют общих внутренних точек. Пишем $F = F_1 + F_2 + \dots$ Под многогранной фигурой будем здесь понимать многогранное тело (см. § 4 гл. II ч. 2) или, проще говоря, объединение конечного числа тетраэдров (симплексов).

Понятие объема определяется так же, как понятие площади, с той лишь разницей, что на место многоугольных фигур ставятся многогранные фигуры; при этом все выводы, какие были сделаны для площади, переносятся на объем.

Объемом многогранной фигуры называется величина со свойствами инвариантности и аддитивности, т. е.: 1) равные фигуры имеют один и тот же объем; 2) если фигура составлена из нескольких многогранных фигур, то ее объем равен сумме их объемов.

Соответственно определяется *численный объем* $V(F)$ при данной единичной фигуре E. Это такое число, что: 1) $V(F) > 0$; 2) если $F \cong F'$, то $V(F) = V(F')$; 3) $V(F_1 + F_2) = V(F_1) + V(F_2)$; 4) $V(E) = 1$. За E принимают единичный куб.

Аналогично теореме I § 1 о площади выполняется

Теорема 1. *При заданной единичной фигуре E каждой многогранной фигуре отвечает, и притом единственное, число со свойствами* 1–4. *Если фигуру E заменить на E', то все численные объемы умножатся на один и тот же множитель:*

$$V'(F) = kV(F); \quad k = V'(E) = \frac{1}{V(E')}.$$

Для любых фигур принимают определение.

Объемом фигуры F называется величина, которая не больше объемов многогранных фигур, содержащих F, и не меньше объемов многогранных фигур, содержащихся в F, при условии, что разности этих объемов могут быть сколь угодно малыми.

Тогда говорим, что фигура имеет определенный объем; иначе у нее нет определенного объема.

Теорема 2. *Фигура имеет определенный объем тогда и только тогда, когда объем ее границы равен нулю. Для таких фигур объем обладает всеми свойствами объемов многогранных фигур, указанными в теореме* 1.

Доказательство теорем 1, 2 может быть построено совершенно так же, как доказательство соответствующих теорем для площади, исходя из определения объемов по способу его измерения. Именно, объем можно определить как величину, измеряемую объемом кубов решетки, заключающихся в фигуре и покрывающих фигуру, если у этих чисел есть общий предел.

Пространство разбивается на единичные кубы E_1, примыкающие друг к другу целыми гранями, так что они образуют кубическую решетку. Этим кубам приписывается значение $V(E_1) = 1$. Эти кубы делятся на равные кубы E_2, и получается вторая решетка; если при этом каждый куб E_1 делится на N_2 кубов E_2, то этим кубам приписывается объем $V(E_2) = 1/N_2$. Затем кубы E_2 делятся на равные и т. д. Получаем последовательность неограниченно измельчающихся кубических решеток.

Объем фигуры определяется с помощью этих решеток буквально так же, как определялась площадь с помощью квадратных клеток.

Последующие доказательства свойств объема происходят почти дословно так же, как для площади, лишь с тремя отличиями: 1) вместо квадратов нужно говорить о кубах;

2) вместо доказательства того, что площадь отрезка равна нулю, нужно доказывать, что объем квадрата равен нулю (поскольку любой многоугольник можно заключить в квадрат);

3) вместо отражения в прямой нужно при доказательстве инвариантности объема воспользоваться отражением в плоскости.

В остальном все выводы проходят так же, как для площади. Они будут те же и в n-мерном пространстве — для n-мерного объема. Тогда в пункте 2) нужно рассматривать не квадрат, а $(n-1)$-мерный куб.

Читатель сможет сам проделать все эти выводы с немалой для себя пользой.

Достоинство проведенного выше доказательства существования площади состоит, между прочим, как раз в том, что оно без изменений переносится на объем любого числа измерений.

Глава III

ДРУГИЕ ОСНОВАНИЯ ГЕОМЕТРИИ

§ 1. Координаты

Наши аксиомы обеспечивают возможность ввести прямоугольные координаты на плоскости и в пространстве и тем опереть аналитическую геометрию на аксиоматическое основание. В том, как при этом вводятся координаты, нет ничего нового в сравнении с тем, как это делается в аналитической геометрии; новое — в том, что нужно проследить, как это основывается на аксиомах и выводах из них, сделанных в гл. I.

Мы начинаем с того, что выбираем какой-либо отрезок и принимаем его за масштаб длин. Тогда согласно выводам § 3 гл. I, каждому отрезку относится численная длина с известными свойствами, указанными там же, и для каждого действительного числа x существуют отрезки с численной длиной x.

Введем координаты на прямой. Для этого докажем:

Каждая точка O любой прямой a делит ее (за вычетом самой точки O) на две полупрямые, т. е. на две такие фигуры, что отрезок, соединяющий точки одной полупрямой, не содержит точку O, а соединяющий точки из разных полупрямых — содержит точку O *внутри* себя (т. е. точка O лежит на нем).

Доказательство. Пусть дана прямая a и на ней точка O (рис. 28). Возьмем на a еще точку A и, продолжив отрезок AO за точку O, получим отрезок AB, на котором лежит O. Прямая a представляет собой объединение отрезков MN, содержащих точки A, B. И так как точка O лежит на отрезке AB, то она лежит на каждом из отрезков MN (теорема 2, § 1 гл. 1) и делит их на отрезки OM, ON (по аксиомам $\mathrm{II}_{2,3}$); считаем, что $A \in OM$, $B \in ON$. Объединения этих отрезков за вычетом точки O и дают полупрямые, на которые точка O делит прямую a. Действительно, если, например, точки P, Q лежат на OM, то отрезок PQ не содержит O; напротив, если P на OM и Q на ON, то O на PQ (по теореме 5, § 1 гл. I).

Рис. 28

Теперь вводим координаты на прямой. На данной прямой берем какую-либо точку O; она делит прямую на две полупрямые; одну из них

a^+ называем *положительной*, другую a^- — *отрицательной*. Точке M на прямой a относим число x_M — «координату точки M» — по известному правилу: 1) если $M \in a^+$, то $x_M = |OM|$, т. е. равно численной длине отрезка OM; 2) если $M \in a^-$, то $x_M = -|OM|$; 3) если $M = O$, то $x_M = 0$.

Так между точками прямой и вещественными числами устанавливается взаимно однозначное соответствие. При этом отрезкам соответствуют числовые промежутки (замкнутые интервалы): точкам отрезка MN отвечают — если $x_M < x_N$ — числа $x \in [x_M, x_N]$, т. е. $x_M \leqslant x \leqslant x_N$. При этом $|MN| = x_N - x_M$.

Эти известные утверждения доказываются на основе теорем из «алгебры отрезков» и доказанной здесь теоремы о разбиении прямой на полупрямые. Проведите эти доказательства (например, если $M \in a^-$, $N \in a^+$, то O на MN по доказанной теореме; $|MN| = |MO| + |ON|$ по аксиоме сложения; так как $M \in a^+$, $N \in a^-$, то по определению $x_M = -|OM|$, $x = |ON|$. Поэтому $|MN| = x_N - x_M$).

Координаты на плоскости вводят известным путем как в аналитической геометрии, используя проведение перпендикуляров и аксиому параллельных отрезков (проделайте соответствующие выводы с должными ссылками). Координаты точки M — это координаты x, y ее проекций M_x, M_y на взаимно перпендикулярные оси x, y (рис. 29).

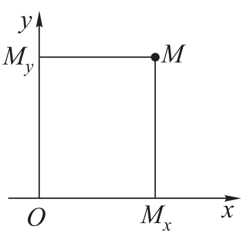


Рис. 29

Таким образом, между точками плоскости и парами координат устанавливается взаимно однозначное соответствие. Точка представляется своими координатами x, y.

Длина отрезка AB выражается через координаты его концов известной формулой

$$|AB| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}. \tag{1}$$

Эта формула выводится, как это сделано еще в § 3 гл. I ч. 1, из теоремы Пифагора, доказательство которой указано в § 2 гл. II ч. 2.

Координаты в пространстве. В пространстве есть, по крайней мере, 4 точки, не лежащие в одной плоскости (по аксиоме IX_1 § 6 гл. I). Берем какую-нибудь точку O и еще две точки A, B. По аксиоме VIII_1 через них проходит плоскость. В этой плоскости α проводим через точку O две взаимно перпендикулярные прямые a, b и вводим координаты x, y, как описано выше. Прямые a, b становятся осями x и y.

Далее, проводим через точку O прямую c, перпендикулярную прямым a и b, и вводим на ней координату z с тем же началом O. Прямая c становится осью z.

Теперь каждой точке M сопоставляются в качестве ее координат x_M, y_M, z_M координаты ее проекций M_x, M_y, M_z на оси x, y, z.

Однако нужно обосновать последние построения, начиная с проведения прямой c. Сделаем это (рис. 30).

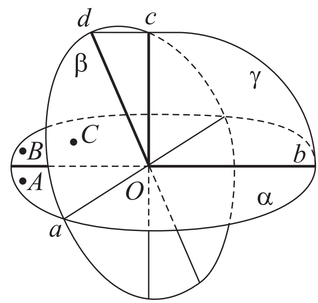


Рис. 30

По аксиоме IX_1 существует точка C, не лежащая в плоскости α. Через нее и прямую a проходит плоскость β. В этой плоскости проводим через точку O прямую d, перпендикулярную прямой a. Через прямые b, d проводим плоскость γ.

Так как прямые b, d перпендикулярны a, то всякая прямая, проходящая в плоскости γ через точку O, перпендикулярна a.[36] Поэтому прямая c, проходящая в этой плоскости перпендикулярно прямой b, будет перпендикулярна и a, и b.

Итак, каждой точке M сопоставляется тройка (упорядоченная тройка) вещественных чисел x_M, y_M, z_M — координаты ее проекций M_x, M_y, M_z на оси x, y, z.

При этом оказывается выполненным и обратное: каждой тройке чисел x_0, y_0, z_0 соответствует определенная (единственная) точка M с такими координатами: $x_M = x_0$, $y_M = y_0$, $z_M = z_0$.

Пусть даны три вещественных числа x_0 y_0, z_0; найдем точку с такими координатами.

Если одно из чисел x_0, y_0, z_0 равно нулю, скажем, $z_0 = 0$, то берем точку M на плоскости $z_0 = 0$, т. е. на плоскости, содержащей оси x, y с координатами $x_M = x_0$, $y_M = y_0$.

Если все числа x_0, y_0, z_0 отличны от нуля, то берем на плоскости, содержащей оси x, y, точку N с координатами $x = x_0$, $y = y_0$ (по доказанному о координатах на плоскости такая точка есть, и притом единственная). Через эту точку и ось z проводим плоскость α. На оси z берем точку M_z с координатой $z = z_0$. В плоскости α проводим отрезок $NM \perp ON$, равный OM_z и расположенный с той же стороны, что и

[36] Этот известный факт мы здесь не доказываем. Доказательство в понятиях элементарной геометрии без векторов можно найти в школьных учебниках, например, в книге: *Александров А. Д., Вернер А. Л., Рыжик В. И.* Геометрия 9–10. М.: Просвещение, 1984.

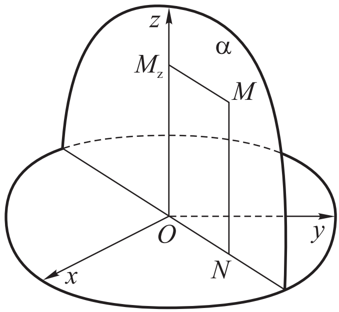


Рис. 31

OM_z. Тогда $MM_z \perp OM_z$, так что M_z есть проекция точки M на ось z, и, стало быть, ее координата $z_M = z_0$ (рис. 31).

Проекции точки M на оси x, y совпадают с проекциями точки N вследствие известной теоремы «о трех перпендикулярах». В более разумной форме она как раз и утверждает, что если N — проекция точки M на плоскость α и a — прямая в плоскости α, то проекции точек M и N на прямую a совпадают[37].

В общем, введение координат устанавливает взаимно однозначное соответствие между точками пространства и тройками чисел. И так же как в случае плоскости, мы получаем возможность выражать утверждения стереометрии на языке координат — на языке алгебры и анализа. В частности, длина отрезка выражается через координаты его концов известной формулой.

Векторы. На плоскости, как и в пространстве, можно определить векторы и операции с ними, как это сделано в аналитической геометрии, но строго опираясь на каждом шагу на аксиомы и их ближайшие следствия. Пользуясь векторами, можно затем ввести координаты. Проследить такое аксиоматическое обоснование векторного исчисления — очень полезная задача.

§ 2. Аналитические основания геометрии

В предыдущем параграфе было показано, что введение прямоугольных координат обеспечено аксиомами. В частности, для плоскости это означает, что из принятых нами аксиом планиметрии вытекает: существует такое взаимно однозначное соответствие между точками и парами чисел, при котором численная длина отрезка выражается известной формулой (1) § 1.

Оказывается, это можно положить в основание планиметрии вместо наших аксиом и дать таким образом чисто аналитическое ее обос-

---

[37] Доказательство. Точка N — проекция M на α, и пусть $P \in a$. Треугольник MNP прямоугольный; угол N прямой. Поэтому $|MP|^2 = |MN|^2 + |NP|^2$. Отсюда следует, что $|MP|$ достигнет минимума вместе с $|NP|$. А минимальное расстояние — по перпендикуляру. Следовательно, одновременно $MP \perp a$ и $a \perp NP$.

нование в отличие от геометрического. Другими словами, можно не вводить координаты и не начинать аналитическую геометрию, исходя из элементарной планиметрии, а наоборот: приняв за основу начала аналитической геометрии, выводить из них планиметрию — ее геометрические аксиомы и теоремы. То же можно сделать для стереометрии, но мы ограничимся планиметрией, ради простоты.

Итак, следуя сказанному, мы принимаем следующие исходные положения:

1. *Точки плоскости поставлены во взаимно однозначное соответствие с парами чисел* (x, y) (*со всевозможными парами*).

2. *Каждой паре точек* A, B *сопоставляется число — «численное расстояние между ними»*:

$$|AB| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}. \tag{1}$$

3. *Геометрический смысл имеют те и только те соотношения, которые определяются через эти расстояния, и притом так, что они — эти соотношения — сохраняются при умножении всех расстояний на любой, но один и тот же положительный множитель.*

Смысл этого последнего условия ясен: численное расстояние появляется, когда выбран масштаб, а при изменении масштаба все расстояния умножаются на одно и то же число. Геометрические факты не зависят от произвольно выбранного масштаба, поэтому они сохраняются при умножении всех расстояний на одно и то же положительное число. Кроме того, тут же заключено условие, что сами по себе значения координат — числа x, y — не должны входить в геометрические утверждения, потому что, координаты имеют геометрический смысл не сами по себе, но только в отношении к данным осям координат. А по отношению к осям они определяются также через расстояния. Например, уравнение эллипса выражает не свойства эллипса самого по себе, а его свойства в отношении к осям тех координат, в которых записано это уравнение. (Различие взглядов — геометрического и аналитического (алгебраического) мы уже обсуждали в ч. I, в главах, посвященных аналитической геометрии. Теперь его важно вспомнить, чтобы вводя аналитические — координатные — основания геометрии, не потерять в них ее специфику.)

Приняв три сформулированные положения, можно дать соответствующее определение планиметрии:

Планиметрия — это теория, основанная на трех высказанных положениях, принятых в качестве аксиом. Эти аксиомы несколько

необычны по содержанию и формулировкам, но раз мы их принимаем как основание теории, то они — ее аксиомы. Можно сказать, они дают координатное, аналитическое основание планиметрии. Впрочем, нужно еще доказать, что эти координатные аксиомы — сокращенно КА — действительно дают основание планиметрии — той же, какая основана на геометрических аксиомах — сокращенно ГА. Должна быть доказана следующая

Теорема 1. *Обе системы аксиом* КА *и* ГА *равносильны, т. е.:* 1) *при геометрической аксиоматике выполняются три координатные аксиомы и* 2) *обратно: при координатных аксиомах выполняются геометрические аксиомы, если соответствующим образом определены их основные понятия.*

Доказательство. Установим первое. Мы доказали в предыдущем параграфе, что при геометрической аксиоматике две первые координатные аксиомы выполняются. Нужно доказать, что третья тоже выполняется, т. е. нужно доказать такое утверждение:

Все основные понятия геометрической аксиоматики определяются через расстояния (1), *взятые с точностью до множителя.*

1. Точки приняты без определения в ГА, как и в КА, а потому не нуждаются в определении.

2. Точка M лежит на отрезке AB. Это равносильно тому, что

$$|AM| + |BM| = |AB| \tag{2}$$

и M отлична от A и B.

3. Отрезок AB — множество точек M с условием (2) плюс сами точки A и B.

4. Точки A, B являются концами отрезка AB, как он только что определен.

5. Отрезки AB, CD равны, если $|AB| = |CD|$ (это равенство, как и (2), сохраняется при умножении расстояний на один и тот же множитель).

Итак, все основные понятия ГА определяются через численные расстояния, взятые с точностью до множителя. В результате все три аксиомы КА выполняются в ГА — в геометрической планиметрии. Нужно доказать обратное, а именно:

При указанной только что интерпретации основных понятий ГА *все ее аксиомы выполняются в* КА — *в координатной планиметрии.*

Д о к а з а т е л ь с т в о. Сделаем прежде всего очевидное, но важное замечание: соотношения, определяемые расстояниями, сохраняются при преобразованиях, сохраняющих расстояние.

Такими преобразованиями являются:

$$\text{(I)}\ x' = x + x_0, \quad y' = y + y_0 \qquad \text{— «перенос»;}$$

$$\text{(II)}\ \begin{aligned} x' &= ax + by, \\ y' &= -bx + ay, \end{aligned} \qquad a^2 + b^2 = 1 \quad \text{— «поворот»;}$$

$$\text{(III)}\ x' = x,\ y' = -y \qquad \text{— «отражение в прямой».}$$

Непосредственно проверяется, что при таких преобразованиях расстояние сохраняется.

Для (I), (III) это очевидно. Проверим для (II):

$$\begin{aligned}
&(x_1' - x_2')^2 + (y_1' - y_2')^2 = \\
&\quad = [a(x_1 - x_2) + b(y_1 - y_2)]^2 + [-b(x_1 - x_2) + a(y_1 - y_2)]^2 = \\
&\quad = a^2(x_1 - x_2)^2 + b^2(y_1 - y_2)^2 + b^2(x_1 - x_2)^2 + a^2(y_1 - y_2)^2 = \\
&\qquad\qquad = (x_1 - x_2)^2 + (y_1 - y_2)^2.
\end{aligned}$$

Теперь назовем *лучом* OA, идущим из начальной точки $O(x_0, y_0)$ через точку $A(x_A, y_A)$, множество точек M с координатами

$$\begin{aligned} x &= x_0 + t(x_A - x_0), \\ y &= y_0 + t(y_A - y_0), \end{aligned} \qquad t \in [0, \infty).$$

В частности, луч, идущий из точки $(0, 0)$ через точку A, задается формулами $x = tx_A$, $y = ty_A$.

Прямой, проходящей через точки A, B назовем множество точек с координатами, выражаемыми такими же формулами, но при t, пробегающем всю числовую прямую от $-\infty$ до $+\infty$.

Всякий луч можно преобразованиями (I), (II) перевести в луч, идущий из точки $(0, 0)$ по точкам $(x, 0)$.

Действительно, во-первых, очевидно, что всякий луч можно преобразованием (I) перевести в луч, исходящий из точки $(0, 0)$. (Если начало данного луча — это точка (x_A, y_A), то производим преобразование $x' = x - x_A$, $y' = y - y_A$.)

Теперь мы имеем луч с началом в точке $(0, 0)$. Пусть он проходит через точку (x_A, y_A), так что он образован точками (tx_A, ty_A). Произведем преобразование (II), положив

$$a = \frac{x_A}{r_A}, \quad b = \frac{y_A}{r_A}, \quad r_A = \sqrt{x_A^2 + y_A^2}.$$

Это допустимо, так как здесь $a^2 + b^2 = 1$. Любая точка луча (tx_A, ty_A) отобразится в точку с координатами

$$\begin{aligned} x' &= \frac{x_A}{r_A}tx_A + \frac{y_A}{r_A}ty_A = tr_A, \\ y' &= -\frac{y_A}{r_A}tx_A + \frac{x_A}{r_A}ty_A = 0. \end{aligned} \tag{3}$$

Итак, точки луча отображаются в точки $(tr_A, 0)$, т. е. луч отображается на луч, идущий из $(0, 0)$ через $(r_A, 0)$. Одновременно прямая AB отображается на прямую $y = 0$. Достаточно брать любые t в (3), чтобы убедиться в этом.

Теперь обратимся непосредственно к доказательству теоремы 1. Докажем, что *в* КА *выполняются все линейные аксиомы из* ГА.

На прямой $y = 0$ между точками и числами x имеется взаимно однозначное соответствие. При этом отрезок AB соответствует числовому промежутку с концами x_A, x_B. Поэтому можно непосредственно проверить, что все линейные аксиомы здесь выполняются (читателю полезно провести эту проверку). В частности, аксиома непрерывности означает здесь следующее. Пусть $x_A = a$, $x_B = b$ и $a < b$. Пусть точки отрезка — промежутка $[a, b]$ — разделены между множествами F, F' так, что для $x \in F$ и $x' \in F'$ всегда $x \leqslant x'$ и F, F' содержат числа, отличные от a и b. Пусть c — точная верхняя граница множества F. Тогда, само собой, при всяком $x \in F$ будет $x \leqslant c$. И при всяком $x' \in F'$ будет $x' \geqslant c$. (Если бы в F было какое-то $x_0' < c$, то нашлось бы $x_1 \in F$ такое, что $x_1 > x_0'$ вопреки определению множеств F и F'.) Итак, при всяких $x \in F$, $x' \in F'$ будет $x \leqslant c \leqslant x'$ или $[ax] \subset [ac] \subset [ax']$. Для отрезков с концами в точках с такими координатами это значит, что $AX \subset AC \subset AX'$ — в согласии с аксиомой непрерывности.

Итак, линейные аксиомы выполняются на прямой $y = 0$. А так как в нее можно перевести всякую прямую, сохраняя расстояние, то и на всякой прямой эти аксиомы выполняются. И так как любой луч можно перевести в луч на прямой $y = 0$ и обратно, то на всяком луче откладывается отрезок, равный данному.

Таким образом, все линейные аксиомы выполняются всюду. □

Теперь проверим, что выполняются также все плоскостные аксиомы.

1. Аксиома деления плоскости. Прямая $y = 0$ разбивает все не лежащие на ней точки на два класса: у одних $y > 0$, у других $y < 0$. Отрезок, соединяющий точки разных классов, содержит точку с $y = 0$. На отрезке же, соединяющем точки одного класса, y сохраняет знак. Таким образом, утверждение аксиомы деления плоскости выполняется для прямой $y = 0$. Любую прямую можно перевести в прямую $y = 0$ преобразованием всей плоскости, сохраняющим расстояние. Поэтому и для всякой прямой утверждение аксиомы выполняется. □

2. Аксиомы угла. *Углом* будем называть пару лучей с общим началом, не содержащихся в одной прямой (это настоящий угол). Лучи — стороны угла, общее их начало — его вершина; понятие поперечины понимаем как и раньше.

Возьмем какой-нибудь угол ab с вершиной $(0, 0)$ и со стороной a, содержащейся в прямой $y = 0$. Преобразованием плоскости, сохраняющим расстояние, переводим луч a в какой-либо данный луч a_1. Тогда луч b перейдет в некоторый луч b_1. Вследствие сохранения расстояний все соответственные поперечины углов ab и a_1b_1 равны. Значит, $\angle\, ab = \angle\, a_1b_1$.

Возможно, угол a_1b_1 отложен от луча a_1 не по ту сторону, какую мы могли заранее задать. Но тогда сначала заменим угол ab на ab_2, у которого луч b_2 получается из b переменой знака y (т. е. преобразованием (III)). Очевидно, $\angle\, ab_2 = \angle\, ab$. Теперь преобразование, переводящее луч a в a_1, переведет луч b_2 в такой, который лежит от a_1 по другую сторону, чем b_1.

Таким образом, угол, равный данному со стороной на прямой $y = 0$, можно отложить от данного луча по любую сторону от него.

Если теперь дан любой угол ab и любой луч a_1, то переводим угол ab в такой, у которого сторона лежит на прямой $y = 0$. А потом этот угол переводим в угол со стороной a_1.

Таким образом, для всякого угла ab есть при любом данном луче a_1 два угла со стороной a_1, равных $\angle\, ab$: один по одну сторону от a_1, другой — по другую.

Но этим еще аксиомы угла не доказаны: нужно еще доказать 1) единственность отложенного угла, 2) равенство у двух углов всех соответственных поперечин, когда есть хотя бы одна пара равных. Для того чтобы установить это, докажем:

При любых двух данных различных точках A, B существует не

более двух точек C, для которых расстояния $|AC|$, $|BC|$ имеют заданные значения; если есть две такие точки, то они лежат по разные стороны от отрезка AB.

Пусть сначала точки A и B — это $(0, 0)$ и $(b, 0)$, $b \neq 0$. Пусть r_1, r_2 — данные расстояния от A и B, так что если x, y — координаты точки C,

$$r_1^2 = x^2 + y^2,$$
$$r_2^2 = (x - b)^2 + y^2 = x^2 + y^2 - 2bx + b^2.$$

Отсюда $r_2^2 = r_1^2 - 2bx + b^2$. Из этого равенства находим x, и тогда из первого равенства получаем не более двух возможных значений y. Два значения y различаются только знаком, т. е. соответствующие точки C_1, C_2 лежат по разные стороны от прямой $y = 0$ (одно значение будет, когда $y = 0$, — точка C лежит на прямой $y = 0$).

Этот вывод мы получили, когда точки A, B лежат на прямой $y = 0$. Если теперь точки A, B любые, то мы можем перевести прямую AB в прямую $y = 0$ преобразованием всей плоскости, сохраняющей все расстояния. Точки A, B перейдут в некоторые A_0, B_0 с $|A_0B_0| = |AB|$, и для них найдется не более двух точек C_0 с данными расстояниями $|A_0C_0| = r_1$, $|B_0C_0| = r_2$. Ввиду сохранения расстояний есть не более двух точек C на данных расстояниях от A и B; и если их две, то они лежат по разные стороны от прямой AB. $\square$

Пусть теперь дан угол ab и от луча a_1 по данную от него сторону отложен угол a_1b_1. Допустим, есть еще другой угол a_1b_2, равный ab, по ту же сторону от луча a_1. Тогда у углов a_1b_1 и a_1b_2 есть соответственные поперечины, которые равны, но концы на лучах b_1 и b_2 у них разные. Если O — начало луча a_1, то это будут какие-то поперечины, AB_1, AB_2 с $|OB_1| = |OB_2|$ и $|AB_1| = |AB_2|$. Но, по доказанному, с одной стороны от AB_1 может быть только одна точка с данными расстояниями от O и A. Следовательно, предположенное невозможно. И, стало быть, угол a_1b_1, равный ab, только один.

Теперь докажем:

Если у двух углов есть пара равных соответственных поперечин, то все их соответственные поперечины равны, т. е. углы равны.

Допустим, у углов ab и a_1b_1 есть пара равных соответственных поперечин AB, A_1B_1, так что $|OA| = |O_1A_1|$, $|OB| = |O_1B_1|$, $|AB| = |A_1B_1|$. По доказанному, от луча a_1 можно отложить угол a_1b_2, равный углу ab, с той же стороны, где лежит угол a_1b_1. И так как $\angle a_1b_1 = \angle ab$, то $\angle a_1b_2 = \angle a_1b_1$. По равенству этих углов у угла a_1b_2 есть поперечина A_1B_2, соответствующая и равная поперечине A_1B_1 угла a_1b_1. При этом

отрезок O_1A_1 общий у этих углов, так что имеем $|O_1A_1| = |O_1A_1|$, $|O_1B_1| = |O_1B_2|$, $|A_1B_1| = |A_1B_2|$. Но так как точки B_1, B_2 лежат с одной стороны от отрезка O_1A_1, то они совпадают. Тем самым у лучей b_1, b_2 есть общие точки B_1, B_2, и лучи совпадают. Это значит, что угол с одной поперечиной, равной соответственной поперечине угла ab, равен углу ab, — у них все соответственные поперечины равны, что и требовалось доказать. □

Итак, выполнение аксиом угла проверено.

3. Аксиома параллельности. Возьмем точки $A(a, 0)$, $B(b, 0)$ и $C(a, c)$, $D(b, c)$; $a, b, c > 0$. Отрезки AC, BD равны, т. е. $|AC| = |BD| = c$, расположены с одной стороны от AB и $|CD| = |AB|$. Отрезки AC, BD образуют с AB прямые углы, т. е. углы, равные своим смежным. Это проверяется совсем просто, так как по доказанному углы равны, если у них есть пара равных соответственных поперечин.

Таким образом, аксиома параллельных отрезков выполняется, когда отрезок AB лежит на прямой $y = 0$. Но преобразованием, сохраняющим расстояние, можно любой отрезок перевести в отрезок на прямой $y = 0$. Следовательно, аксиома параллельных выполняется при любых отрезках AB. □

На этом проверка аксиом завершается, и таким образом доказано, что в аналитической геометрии, основанной на КА — на координатной аксиоматике — выполняется ГА — геометрическая аксиоматика. А так как обратное уже было установлено, то, стало быть, доказана равносильность обеих аксиоматик, а с ней и теорема 1. □

Аналитическая аксиоматика стереометрии. В основу стереометрии можно положить следующие координатные аксиомы:

1. *Точки пространства поставлены во взаимно однозначное соответствие с тройками чисел x, y, z — со всевозможными упорядоченными тройками.* Эти числа называем *координатами* точки.

2. *Каждой паре точек A, B сопоставлено число* — «***численное расстояние***»

$$|AB| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}.$$

3. *Геометрический смысл имеют те и только те соотношения, которые определяются через эти расстояния, и притом так, что они — эти соотношения — сохраняются при умножении всех расстояний на одно и то же, но любое положительное число.*

Отрезки, отношения точек к отрезкам определяются так же, как и выше — в случае планиметрии; равенство отрезков определяется по

равенству расстояний между концами. Прямую AB можно определить как множество точек с координатами $x = x_A + t(x_B - x_A)$ и аналогично для y и z; $t \in (-\infty, +\infty)$.

Плоскость можно задавать тремя точками A, B, C, не лежащими на одной прямой, как множество точек с координатами

$$x = x_A + t(x_B - x_A) + s(x_C - x_A),$$

и аналогично для y и z; t и s принимают все значения «от $-\infty$ до $+\infty$».

Координатные аксиомы равносильны геометрическим. Доказывать это мы не будем. Принципиально важный факт, что геометрию можно определять аналитически, пользуясь координатами, выяснен на примере планиметрии.

§ 3. Аксиоматика в отвлеченном понимании; ее модель, непротиворечивость, независимость, полнота

Аксиоматику геометрии с основными объектами — отрезками или прямыми — можно понимать двояко: наглядно содержательно и отвлеченно.

В первом случае основные понятия — объекты и отношения — представляются в их исходном наглядном, хотя и идеализированном, смысле, а аксиомы представляют собой описание свойств этих объектов и отношений. Так, например, отрезок мыслится обычно в виде идеально тонкой черты на бумаге или натянутой нити.

В противоположность этому при отвлеченном понимании аксиоматики ее понятия толкуются как относящиеся к объектам и отношениям «произвольной природы», лишь бы для них выполнялось сказанное в аксиомах. При таком понимании отрезок — это просто какой-то мыслимый объект, для которого вместе с другими объектами, называемыми отрезками и точками, с отношениями между ними, названными в аксиомах, выполняется все то, что сказано в аксиомах. Аналогичное можно сказать о точках и об основных отношениях.

Эту отвлеченную точку зрения впервые особенно ясно выразил Гильберт, начав свои «Основания геометрии» следующими примечательными словами:

«Мы мыслим три различных системы вещей: вещи первой системы мы называем точками и обозначаем A, B, C, …; вещи второй системы мы называем прямыми и обозначаем a, b, c, …; вещи третьей системы

мы называем плоскостями и обозначаем α, β, γ, ... Мы мыслим точки, прямые и плоскости в определенных соотношениях и обозначаем эти соотношения различными словами, как-то: «лежать», «между», «конгруэнтный»... Точное и для математических целей полное описание этих соотношений достигается аксиомами геометрии»[38].

Таким образом, точки, прямые, плоскости — это какие-то мыслимые вещи или, как более принято у нас говорить, объекты, находящиеся в некоторых отношениях, все необходимые для геометрии свойства которых описываются аксиомами. Как бы подчеркивая это отвлечение от наглядных представлений, Гильберт говорит: «точка принадлежит прямой» и «прямая принадлежит точке». Первая аксиома у него звучит так: «Для любых двух точек A, B существует прямая a, принадлежащая каждой из этих двух точек A, B». Ни прямая, ни плоскость не мыслятся «состоящими из точек». Отвлечение от наглядности в исходных формулировках обеспечивает уверенность, что все необходимое для логического построения геометрии действительно выражено в аксиомах.

Также с отвлечением от наглядности понимаются при абстрактном взгляде основные объекты и отношения нашей, да и всякой другой аксиоматики. Поэтому на вопрос, что такое прямая в системе Гильберта, или отрезок — в нашей аксиоматике, так же как отношение «лежит на» и т. п., следует ответ: это то, о чем говорится в аксиомах.

Таким образом, в отвлеченном понимании аксиомы представляют собой определение основных объектов и отношений. *Отрезком называется то, что под таким названием фигурирует в аксиомах.* И так же можно определить другие основные понятия.

Такие определения называются *аксиоматическими.* Обычное определение понятия состоит в том, что оно разъясняется через другие. В отличие от этого в аксиоматическом определении понятия разъясняются совместно, через их связь, как это видно на примере определения отрезка.

Отвлеченно рассматриваемая аксиоматика сама по себе ни к чему сколько-нибудь определенному не относится, так что не ясно, какой смысл она имеет; не представляет ли она просто набор слов? Для аксиоматики евклидовой геометрии это не так, потому что в ней, как бы отвлеченно мы ее ни рассматривали, заключен первоначальный реальный смысл. Но вот когда Лобачевский создавал свою геометрию, смысл ее был неясен. На вопрос о ее предмете можно было ответить

[38] *Гильберт Д.* Основания геометрии. М.; Л.: ОГИЗ, 1948. С. 56.

только, что это то, о чем мы рассуждаем. Лобачевский называл свою геометрию «воображаемой». Предмет, к которому она могла относиться, был открыт позже.

Вообще, для того чтобы отвлеченная аксиоматика получила более определенный смысл, нужно найти предмет, к которому она могла бы относиться, модель, где бы аксиоматика выполнялась, относясь к определенным объектам и отношениям.

Модель или, как еще говорят, *интерпретация аксиоматики представляет собой, коротко говоря, совокупность некоторых объектов с отношениями, для которых выполняются аксиомы.*

Говоря подробнее, основные объекты и отношения аксиоматики сопоставляются объектам модели и определенным имеющимся между ними отношениям. Если при этом для объектов и отношений модели выполняется то, что говорится в аксиомах (если эти объекты и отношения назвать словами, принятыми в аксиомах), то мы и получаем тем самым интерпретацию аксиоматики. Или, как еще говорят, аксиомы реализуются на данной модели. При этом имеют в виду математическую модель. Ссылка на то, что евклидова геометрия имеет реальный смысл, обнаруживающийся на опыте, для чистой математики не годится, потому что опыт не может быть математически точным.

Для планиметрии у нас уже есть готовая модель — ее дает координатная аксиоматика. В самом деле, в ней точкам сопоставляются пары чисел, и все дальнейшие основные понятия получают представление в этих числах, и все аксиомы выполняются, как показано в § 2. При этом не нужно относить координаты к каким-то точкам, можно просто считать, что точка в модели и есть пара чисел. Тогда всякая неопределенность в понятии точки исчезает, и мы получаем числовую модель планиметрии.

Непротиворечивость, независимость, полнота. Для отвлеченной аксиоматики самой по себе неизвестно, не могут ли выводы из нее привести к противоречию, когда в одном выводе что-то утверждается, а в другом оно же отрицается.

Такая аксиоматика заведомо не может быть реализована и не имеет смысла. Если же такие противоречия не могут получиться, аксиоматика называется непротиворечивой.

Вообще, рассматривают три существенных свойства аксиом: непротиворечивость, независимость, полноту.

1. Система аксиом называется, как уже сказано, *непротиворечивой*, если из нее не следует какое-либо утверждение вместе с его же отрицанием.

В следующих двух определениях имеется в виду непротиворечивая система аксиом.

2. Аксиомы называются *независимыми*, если ни одна из них не следует из других.

3. Система аксиом называется *полной*, если она не может быть пополнена, т. е. если к аксиомам нельзя добавить никакой аксиомы, которая из них не следовала бы и вместе с тем им не противоречила. При этом, конечно, имеется в виду, что добавляемая аксиома касается тех же основных понятий.

В этих определениях речь идет о том, что́ следует из аксиом. Здесь это можно понимать так: утверждение B следует из системы аксиом A, если как только выполняется A, так выполняется B — во всякой модели аксиом A выполняется утверждение B. Если же есть модель аксиом A, где B не выполняется, то, значит, B не следует из A. Поэтому имеют место три утверждения.

1. Система аксиом непротиворечива, если она реализуется в какой-нибудь модели (в модели не может быть противоречия, чтобы B выполнялось и одновременно не выполнялось).

2. В системе аксиом данная аксиома независима от других, если есть модель, где выполняются все аксиомы, кроме этой. Тем самым она, очевидно, не следует из других.

3. Система аксиом полная, если во всех ее моделях выполняется одно и то же, т. е. всякое утверждение, выражаемое в понятиях аксиоматики, верное в одной модели, выполняется также во всякой другой.

Это равносильно полноте в предыдущем определении, что к аксиомам ничего нельзя прибавить. Действительно, если добавляемое утверждение B верно во всякой модели, значит там, где выполняются аксиомы, выполняется также B, т. е. B следует из аксиом. Если же добавляемое утверждение B не выполняется ни в одной модели, то, значит, в них верно его отрицание, так что добавление его ведет к противоречию.

Такие модели, в которых выполняется одно и то же (выражаемое в понятиях аксиоматики), называют *изоморфными*. Это понятие изоморфизма определяется в общем виде следующим образом.

Представим себе две системы объектов с отношениями между ними. Пусть между объектами этих систем, а также между отношениями установлено такое взаимно однозначное соответствие, что соответственные объекты находятся в соответственных отношениях. То есть если объекты a, b, $\ldots$ одной системы S находятся в отношении R, то соответствующие им объекты a', b', $\ldots$ системы S' нахо-

дятся в соответствующем отношении R', и обратно: если какие-то объекты системы S' находятся в отношении R', то соответствующие им объекты системы S находятся в соответствующем отношении. Такое соответствие называется *изоморфизмом.* А системы, для которых может быть установлено такое соответствие, называются *изоморфными.*

Сказанное выше о полноте системы аксиом выражают так: *система аксиом полная, если все ее модели* (*интерпретации*) *изоморфны.*

Из определения изоморфизма ясно, что если в одной системе S выполняется некоторое утверждение, то соответствующее утверждение выполняется также в изоморфной системе S. Просто потому, что всякое утверждение касается объектов и их отношений. Если назвать соответственные объекты и отношения одними и теми же словами, то все, что относится к одной системе, будет буквально относиться и к другой. Они представляются тогда совершенно одинаковыми. Так и модели полной системы аксиом одинаковы. Но, конечно, только в отношении свойств, относящихся к аксиоматике, так как в моделях могут быть и другие свойства (как, например, в указанной выше модели, когда точка есть пара чисел, точки различаются — скажем, (0, 0) и (1, 1), — но это в нашей аксиоматике ничему не соответствует). В этом же смысле можно сказать, что полная аксиоматика определяет один определенный предмет или, как уточняют, один с точностью до изоморфизма.

Мы представляем себе плоскость как вполне определенный «предмет», так что в планиметрии какое-либо утверждение либо верно, либо неверно — геометрический факт имеет место или его нет. Поэтому аксиоматика планиметрии должна быть полной. Наша аксиоматика полная, как мы сейчас докажем.

Вообще три указанные свойства аксиоматики играют существенно разную роль. Непротиворечивость обязательна, поскольку противоречивая система просто не имеет смысла. Независимость аксиом необязательна, но если ее нет, т. е. есть аксиома, которая следует из других, то такая аксиома лишняя и ее можно исключить. Желательно не иметь таких лишних аксиом, хотя их можно вводить для того, чтобы упростить доказательства теорем. Так мы могли бы, например, включить некоторые теоремы § 3 гл. I в аксиомы.

Полнота, как было объяснено, необходима для аксиом планиметрии. Но вообще в математике большое значение имеют как раз неполные аксиоматики. Такая аксиоматика определяет не один предмет теории, а целый класс их как группы и кольца в алгебре (поскольку

есть, например, неизоморфные группы) или разные типы пространств в обобщенной геометрии.

Выводы о нашей системе аксиом геометрии. Как уже было отмечено, у нас есть числовая модель планиметрии. Точками в этой модели служат пары чисел (x, y), а прочие основные понятия определяются через эти числа — как это сделано в § 2. Наличие такой модели дает основание утверждать, что

Наша система аксиом планиметрии непротиворечива,

или, более точно,

Наша аксиоматика планиметрии непротиворечива, насколько непротиворечива теория вещественных чисел.

Действительно, если бы из наших аксиом можно было бы вывести противоречие, то оно обнаруживалось бы в модели, и тем самым — в теории вещественных чисел, на основе которой модель построена.

Тот же вывод о непротиворечивости верен и для аксиом стереометрии, поскольку для нее тоже есть числовая модель. Мы только не провели доказательства того, что в модели, указанной в конце § 2, действительно выполняются все аксиомы стереометрии.

Уточняя, мы сказали, что наша система аксиом настолько непротиворечива, насколько непротиворечива теория вещественных чисел. Именно это и следует из нашей числовой модели. А возможно ли противоречие в этой модели — это уже другой вопрос — вопрос о непротиворечивости теории вещественных чисел. Есть основание считать ее свободной от возможного противоречия. Но независимо от возможных сомнений в непротиворечивости вещественных чисел сделанный вывод о системе аксиом геометрии имеет важный смысл.

Вещественные числа служат основным предметом всей математики, через них определяются функции и комплексные числа и функциональные пространства и т. д. Поэтому, давая числовую модель аксиом геометрии, мы тем самым логически включаем геометрию в целостную систему математики. И непротиворечивость оснований геометрии опирается таким образом на широчайшее основание в математике. Что же до окончательного доказательства непротиворечивости, то оно не может быть дано: доказывая, мы чем-то пользуемся, и об этом тоже можно спросить: а не скрыто ли здесь возможное противоречие?

Обратимся к вопросу о полноте нашей системы аксиом. Тут мы можем утверждать:

Наша аксиоматика планиметрии полная.

Доказательство. Исходя из аксиом, мы ввели на плоскости прямоугольные координаты и вывели формулу для расстояния между

точками. Стало быть, это можно сделать во всякой модели, где выполняются аксиомы. Значит, всякая модель изоморфна числовой модели, в которой точки — это пары чисел и расстояние задается той же формулой. Таким образом, наша аксиоматика планиметрии полная.

То же верно для аксиоматики стереометрии, поскольку и в пространстве тоже вводятся координаты и все геометрические соотношения могут быть в них выражены.

О независимости наших аксиом планиметрии можно сказать, что есть независимые аксиомы, но есть и зависимые.

Аксиома параллельных отрезков независима; это будет доказано потом.

Аксиома непрерывности независима (это доказывается числовой моделью, в которой, однако, фигурируют не все вещественные числа x, y, а только такие, которые получаются из 1 применением действий типа $a \pm b$, ab, a/b, $\sqrt{a^2+b^2}$).

В аксиоме откладывания угла требуется, что откладывается только один угол, равный данному. Это можно исключить: эта часть аксиомы зависима. Мы ввели ее для простоты, поскольку ее доказательство достаточно сложно.

Выяснение независимости всех аксиом представляет трудность.

Экзотическая модель планиметрии. Представим себе плоскость с прямоугольными координатами x, y. Но будем считать отрезком, соединяющим точки $A(x_0, y_0)$, $B(x_1, y_1)$, кривую, образуемую точками $M(x, y)$ с координатами

$$\begin{aligned} x &= \sqrt[3]{(1-t)x_0^3 + tx_1^3}, \\ y &= (1-t)y_0 + ty_1, \quad 0 \leqslant t \leqslant 1. \end{aligned} \tag{1}$$

Припишем «отрезку» AB «длину»

$$|AB| = \sqrt{(x_1^3 - x_0^3)^2 + (y_1 - y_0)^2} \tag{2}$$

и будем считать равными отрезки равной «длины». Все аксиомы будут выполняться.

Если $x_0 \neq x_1$, то, исключая из (1) параметр t, получим уравнение вида

$$y = kx^3 + 1. \tag{3}$$

А при $x_0 = x_1$ будет $x = x_0$, и при $y_0 = y_1$ будет $y = y_0$, т. е. «отрезки» — это дуги кубических парабол (3) и отрезки прямых $x = \text{const}$, $y = \text{const}$.

Общий прием получения таких моделей состоит в следующем. Полагаем $x = f(\overline{x})$, $y = g(\overline{y})$, где функции f, g обратимы. Отрезок $x = (1-t)x_0 + tx_1$, $y = (1-t)y_0 + ty_1$ заменяется кривой:

$$x = f^{-1}((1-t)f(x_0) + tf(x_1)),$$
$$y = g^{-1}((1-t)g(x_0) + tg(x_1)).$$

Расстояние будет

$$|AB| = \sqrt{(f(x_1) - f(x_0))^2 + (g(x_1) - g(x_0))^2}.$$

Вообще, произведем любое взаимно однозначное отображение плоскости на себя. Отрезки перейдут в какие-то фигуры, и, соответственно все, что касалось отрезков, будет касаться этих фигур. Так что получается модель планиметрии с этими фигурами в качестве «отрезков». В приведенном примере мы произвели преобразование, сопоставляя точкам (x, y) точки с координатами

$$\overline{x} = \sqrt[3]{x}, \quad \overline{y} = y.$$

Отрезок

$$x = (1-t)x_0 + tx_1, \quad y = (1-t)y_0 + ty_1$$

перейдет в

$$\overline{x} = \sqrt[3]{(1-t)\overline{x}_0^3 + t\overline{x}_1^3}, \quad \overline{y} = (1-t)\overline{y}_0 + t\overline{y}_1.$$

Можно также взять $\overline{x} = \sqrt[n]{x}$, $\overline{y} = \sqrt[m]{y}$ с любыми нечетными n и m и т. п.

Такие модели служат иллюстрацией к абстрактному взгляду на аксиоматику. «Отрезок» может быть любой кривой — лишь бы формально выполнялись требования аксиом.

§ 4. Разные системы аксиом

Замкнутые и незамкнутые аксиоматики. Системы аксиом, которые кладутся в основание геометрии, как и других теорий, могут различаться в целом ряде отношений. В предыдущем параграфе была представлена система аксиом планиметрии — всего из трех аксиом, — совершенно отличная от той, какую мы приняли с самого начала в гл. I. Главное различие этих двух систем аксиом — геометрической (ГА) и

координатной (КА) — заключается, конечно, не в том, что в одной аксиом довольно много, а в другой — мало. Главное различие в том, что считается заранее известным в одном и в другом случае. В КА используются вещественные числа, а в ГА — ничто не предполагается известным заранее. Аксиомы ГА можно объяснить наглядно человеку, который не знает математики. Но аксиомы КА будут понятны только тому, кто знает теорию вещественных чисел, и уж во всяком случае он должен понимать формулу для расстояния.

Конечно, во всякой аксиоматике что-то подразумевается известным; минимально — это правила грамматики, логика, а также натуральные числа: один, два и т. п. Во всяком случае, без грамматики и логики ничего нельзя сформулировать. Как нельзя, например, высказать аксиому о проведении отрезка без слова «существует» или без равносильного выражения.

Если в аксиоматике используется только такой необходимый минимум, то мы называем ее замкнутой или исчерпывающей. Если же это не так, если в аксиоматике подразумевается известным или само собой понятным что-либо сверх этого, скажем, понятие из другой теории, то такая аксиоматика будет незамкнутой. Наша система аксиом ГА замкнута, система КА — не замкнута.

Раньше, в гл. I ч. 2, мы ввели систему аксиом с теми же основными понятиями, что ГА, и с во многом сходными аксиомами, но включающую аксиомы длины и меры угла. В этих аксиомах используются вещественные числа, и, стало быть, такая система аксиом тоже незамкнута. Ее можно обозначить как НГА — «не вполне геометрическая аксиоматика», поскольку она геометрическая, но отступает от чистой геометрии, включая понятия численной длины и меры угла.

Для того чтобы сделать незамкнутую аксиоматику замкнутой, исчерпывающей, ее нужно дополнить аксиомами, говорящими об используемых в ней понятиях; в принципе это всегда можно сделать. Так, для того чтобы «замкнуть» систему КА, нужно добавить к ней аксиомы вещественных чисел. Этих аксиом, минимум, 18 (восемнадцать!), так что если прибавить их к КА, то получится довольно внушительный список. Немудрено, что аксиом в КА мало: в ней много подразумевается известным!

Простейший случай неисчерпывающей аксиоматики — тот, когда в ней просто пропущены некоторые аксиомы или, более того, — не выражены в аксиомах какие-либо основные понятия, которыми, однако, пользуются в доказательствах. Так обстоит дело в «Началах» Евкли-

да[39] и во многих прежних и современных курсах геометрии, где наряду с аксиомами опираются на очевидность. Но это «не тот случай»; мы обращаем внимание на аксиоматики, использующие негеометрические понятия.

Системы аксиом, использующие вещественные числа и некоторые другие понятия, могут быть очень удобными, как наша НГА; они распространяются и вводятся в школьные учебники. Но надо ясно понимать, что они не являются замкнутыми и, стало быть, сами по себе *не дают исчерпывающих оснований геометрии.*

В аксиоматике сплошь и рядом фигурирует понятие множества. Но оно настолько укоренилось в основах математики, что его появление в аксиоматике можно не считать нарушением ее замкнутости: можно сказать, включающая его аксиоматика замкнута в рамках общей теоретико-множественной точки зрения. Конечно, у теории множеств есть свои аксиомы, и если иметь это в виду, то аксиоматику, использующую понятие множества, нельзя считать замкнутой. Однако аксиоматизированное понятие множества в полном объеме редко бывает нужно, а в элементарной геометрии оно вообще не нужно. Мы вовсе без него обошлись, введя аксиомы фигуры.

Итак, мы фиксируем фундаментальное различие двух типов систем аксиом — замкнутых и незамкнутых.

Различие аксиоматик по выбору основных понятий. За этим следует другое существенное различие — по основным понятиям: какие объекты и отношения принимаются в аксиоматике за основные. Тут между системами ГА и КА опять же есть громадная разница. В КА только один вид основных объектов — точки и никаких основных отношений; все определяется с помощью координат. В ГА, напротив, (в планиметрии) принимаются два вида основных объектов: точки и отрезки, два вида отношений точек и отрезков: точка — конец отрезка или лежит на отрезке, и отношение равенства отрезков.

Различие по основным объектам встречается как между замкнутыми, так и незамкнутыми системами аксиом. Например, у нас в ГА основными объектами являются отрезки, тогда как обычно основны-

[39] Евклид жил в Александрии в III в. до н. э. — в так называемую эпоху эллинизма, наступившую после походов Александра Македонского. В его труде «Начала» были систематизированы и изложены в логической последовательности основы геометрии, а также элементы теории чисел и геометрически изложенной алгебры. Изложению предшествуют определения основных понятий и формулировки основных положений геометрии — постулатов и «аксиом» (теперь это различие не делается). В этом отношении «Начала» положили начало аксиоматике.

ми объектами служат прямые. Соответственно изменяются и основные отношения, а вместе с этим и сами аксиомы. Рассмотрим примеры.

Обратимся к аксиоматике стереометрии; она представлена у нас в пяти вариантах: три незамкнутые в § 8 гл. I ч. 2 и две замкнутые в § 6 гл. I ч. 6. Последний из этих двух вариантов отличается от всех остальных тем, что в нем плоскость исключена из основных объектов. (Пространственные аксиомы в первых вариантах § 8 гл. I ч. 2 и § 6 гл. I ч. 6 одни и те же; разница — в линейных аксиомах; в общем случае они включают понятие численной длины.)

Аксиоматика Гильберта. Из замкнутых геометрических систем аксиом наиболее известна система аксиом Гильберта[40]. В ней основными объектами являются точки, прямые, плоскости, а основными отношениями служат: 1) точка принадлежит прямой, точка принадлежит плоскости; 2) точка лежит между двумя другими (это равносильно тому, что точка C лежит на отрезке с концами A, B); 3) отношение равенства или, как оно называется у Гильберта, конгруэнтности; это, собственно, два отношения: равенство отрезков и равенство углов. При этом под отрезком понимается пара точек A, B вместе со всеми точками, лежащими между ними. Под углом понимается пара лучей с общим началом, не содержащихся в одной прямой, а луч AB определяется как совокупность таких точек M прямой AB, что A не лежит между M и B.

Аксиомы в системе Гильберта делятся на 5 групп: 1) аксиомы принадлежности, 2) аксиомы порядка, касающиеся отношения «между», 3) аксиомы конгруэнтности, 4) аксиома параллельности, 5) аксиомы непрерывности (их две — вместе они равносильны аксиоме непрерывности в нашей аксиоматике).

Можно в нашей аксиоматике исключить отрезки из основных объектов, а вместе с ними — и отношение точек к отрезкам. Вместо этого введем как основное отношение точек: одна точка лежит *между* двумя другими. Далее можно условиться: если точка C лежит между точками A и B, то будем говорить, что точка C «лежит на отрезке AB». Иначе говоря, мы определяем отрезок AB как геометрическое место точек, лежащих между A и B, плюс сами эти точки; при

[40] Давид Гильберт (1862–1943) — немецкий математик. В его работе «Основания геометрии», появившейся в первом варианте в 1899 г., была дана система аксиом геометрии; одна из первых систем и самая совершенная среди тех, где аксиоматические основания геометрии были представлены так, как их с тех пор понимают. Гильберт был крупнейшим математиком своего времени и оказал чрезвычайное влияние на развитие математики в ряде направлений.

этом они называются концами отрезка AB. (Данное определение отрезка согласуется с аксиомой геометрического места из аксиоматики фигуры.)

Например, аксиома деления отрезка выразится с понятием «между» следующим образом.

Если точка C между A и B, то: 1) всякая другая точка M, лежащая между A и B, лежит либо между A и C, либо между B и C; 2) всякая точка M, лежащая между A и C, лежит также между A и B (и то же для точек M между B и C — стоит лишь поменять обозначение A на B).

Читатель может рассмотреть наши аксиомы с точки зрения замены формулировок с отрезками на формулировки с отношением «между». Формулировки эти будут громоздкими. Недаром Гильберт, приняв, как основное, отношение «между», вводит тем не менее и понятие отрезка, и, как основное отношение, — равенство отрезков.

Равенство отрезков и углов устанавливается в практике не само собой, как в аксиомах, а либо наложением, либо измерением (впрочем, измерение тоже состоит в наложении — масштаба на измеряемый предмет). Поэтому естественно принять наложение за основное понятие геометрии.

Другие аксиоматики евклидовой геометрии. Одновременно с появлением системы аксиом Гильберта была дана итальянским математиком М. Пиери система аксиом только с двумя основными понятиями: «точка» и «движение» — то же, что наложение.

Первая аксиома о движении утверждала, что движение сопоставляет точкам точки. Понятия прямой и другие определялись через движение. Однако ограничение только понятиями точки и движения привело к чрезвычайной сложности аксиом, и система Пиери не получила распространения. Потом немецкий математик Ф. Шур дал систему аксиом с теми же основными понятиями, что у Гильберта, кроме конгруэнтности: вместо этого вводится понятие движения, и конгруэнтность определяется как совмещаемость движением (наложением).

Тогда же в 1904 г. независимо от Гильберта В. Ф. Каган[41] дал систему аксиом, в которой расстояние между точками (или, что равносильно, — длина отрезка) определялось как вещественное число. Тем самым эта система была незамкнута. Равенство отрезков определялось равенством их длин. Такая замена геометрического равенства отрез-

[41] Вениамин Федорович Каган (1869–1953) был профессором МГУ (с 1923 г.), создал школу в дифференциальной геометрии.

ков равенством их длин проводилась в учебниках, принятых у нас в школе на протяжении последних 20 лет[42].

Надо, однако, ясно понимать, что отнесение отрезку определенного числа в качестве его длины не соответствует ни практике, ни геометрии, потому что отрезок имеет определенную численную длину не сам по себе, а только при выборе единицы измерения, и при ее замене численная длина изменяется. И вовсе нелепо, когда в аксиомах вводится градусная мера углов, 180° у развернутого угла, как будто если измерять угол другой единицей, то будет уже другая геометрия[43].

В 1918 г. немецкий математик Г. Вейль дал систему аксиом, основанную на векторной алгебре. Вытеснение геометрии алгеброй распространилось и дошло до школьных учебников[44].

Другие особенности аксиоматик. Помимо двух отмеченных фундаментальных различий в системах аксиом — по тому, что предполагается известным, и по основным понятиям, системы аксиом могут различаться и в других отношениях.

1. Самое простое — это различие «по форме», тогда системы аксиом отличаются только формулировками и компоновкой аксиом: какие аксиомы соединяются в одну или разделяются, а также как они распределяются по группам. Так что тут на самом деле не разные, а одна система аксиом, только выраженная несколько разными способами.

Например, нашу аксиому I_1 существования концов отрезка, § 1 гл. 1, можно разделить на три и даже на четыре аксиомы (сделайте это). Можно было бы делить аксиомы на группы несколько иначе, например, относя аксиому деления плоскости к аксиомам связи, поскольку в ней понятие равенства отрезков не участвует.

Разумно сравнивать системы аксиом по числу входящих в них аксиом только в том случае, если разделить аксиомы, допускающие разделение; например, утверждение «у отрезка есть два и только два конца» делится на два: у каждого отрезка есть два конца; у каждого отрезка есть не более двух концов.

2. Более существенным является различие по аксиомам, т. е. такое,

---

[42] *Колмогоров А. Н., Семенович А. Ф., Черкасов Р. С.* Геометрия 6–8 кл. 8-е изд. 1980–1982. (Первоначально длина в этом учебнике определялась как величина.) *Погорелов А. В.* Геометрия 6–10 кл. М.: Просвещение, 1981.

[43] Так же не отвечает ни практике, ни геометрии определение равенства отрезков по равенству численных длин: сантиметровые отрезки на измерительной линейке равны по геометрическому сравнению, а не потому, что им отнесено число 1.

[44] См., например, принятый в недавнее время учебник стереометрии: *Клопский В. М., Скопец З. А., Ягодовский М. И.* Геометрия 9. М.: Просвещение, 1975, где в качестве дополнения излагается аксиоматика Вейля.

когда системы аксиом отличаются хотя бы некоторыми входящими в них аксиомами по содержанию, а не просто по форме выражения. Так, например, получается, когда наша аксиома параллельных отрезков заменяется на аксиому параллельных прямых.

3. Варианты различия «по аксиомам» представляет случай, когда системы аксиом отличаются «силой» условий, в крайнем случае, когда в одной системе есть аксиома, выводимая из других, так что заключенное в ней условие — лишнее, и такую аксиому можно исключить.

Интересный пример дает аксиома параллельных. Вместо нее можно требовать только, что существуют такая прямая a и не лежащая на ней точка A, что через эту точку проходит не более одной прямой, параллельной a. Отсюда, конечно, с помощью других аксиом, уже следует, что то же верно для любой прямой и любой не лежащей на ней точки. Совершенно так же можно ослабить нашу аксиому параллельных отрезков, требуя только, что существуют такие точки A, B, C, D, что отрезки AC, BD перпендикулярны AB и $CD = AB$.

Аксиому деления плоскости можно заменить аксиомой Паша, как это и сделано выше.

Системы аксиом Д. Гильберта, Ф. Шура, Г. Вейля и модифицированная система аксиом А. В. Погорелова изложены в книге: *Александров А. Д.* Основания геометрии. М.: Наука, 1987. Там же дан обстоятельный очерк развития оснований геометрии.

Глава IV

РАЗНЫЕ ГЕОМЕТРИИ

§ 1. Геометрия Лобачевского; ее модели

Геометрия Лобачевского строится на тех же аксиомах, что геометрия Евклида, с единственной заменой аксиомы параллельных на противоположную:

Аксиома Лобачевского. На плоскости для каждой прямой a через каждую не лежащую на a точку проходит по крайней мере две прямых, не пересекающих данную прямую a.[45]

[45] Строго говоря, в изложенном виде аксиома Лобачевского не является логическим отрицанием аксиомы параллельных прямых. Такое отрицание может быть сформулировано следующим образом:

Через некоторую точку проходят две прямые, не пересекающие некоторой третьей прямой.

Поэтому в геометрии Лобачевского выполняются все теоремы евклидовой геометрии — планиметрии и стереометрии, — основанные на аксиомах за вычетом аксиомы параллельных. Но теоремы, связанные с этой аксиомой, заменяются существенно другими, которые на первый взгляд по большей части выглядят очень странно.

Здесь мы укажем модель геометрии Лобачевского на плоскости и этим решим три задачи: 1) докажем непротиворечивость геометрии Лобачевского на плоскости; 2) докажем независимость аксиомы параллельных от других аксиом планиметрии; 3) покажем наглядный смысл фактов геометрии Лобачевского, как они представляются в модели, которая строится на обычной евклидовой плоскости, т. е., иначе говоря, в рамках планиметрии.

Модель Пуанкаре геометрии Лобачевского. (Французский ученый Анри Пуанкаре (1854–1912) — крупнейший математик. Описываемая далее модель была предложена им в 1882 г.) Роль плоскости Лобачевского играет открытая полуплоскость; роль прямых выполняют содержащиеся в ней полуокружности с центрами на ограничивающей ее прямой и лучи, перпендикулярные этой прямой. Роль наложений выполняют композиции инверсий относительно этих полуокружностей и отражений в лучах. Все аксиомы евклидовой геометрии здесь выполняются, кроме аксиомы параллельных (рис. 32, a), тем самым в этой модели выполняется геометрия Лобачевского.

Опишем эту модель более подробно и докажем сказанное. Берем на обычной евклидовой плоскости какую-нибудь прямую p и ограниченную ею открытую полуплоскость P. Прямую p назовем граничной прямой. Полуплоскость P будет играть роль плоскости Лобачевского; мы будем называть ее «плоскостью» в кавычках. Точками в модели будут точки этой «плоскости», т. е. полуплоскости P. За «прямые» в модели принимаем, во-первых, содержащиеся в P полуокружности, центры которых лежат на граничной прямой (рис. 32, a). «Отрезок» AB в модели — это дуга такой полуокружности с концами A, B.

Подчеркнем, что конец «отрезка» не может быть концом полуокружности, представляющей прямую; ее концы исключены вместе с граничной прямой; «плоскость» — это открытая полуплоскость. Точка «прямой» служит общим началом двух «лучей» — двух дуг полуокружности (с исключенными концами). «Углом» назовем фигуру из

С учетом остальных аксиом обе формулировки оказываются равносильными, как это станет ясно дальше.

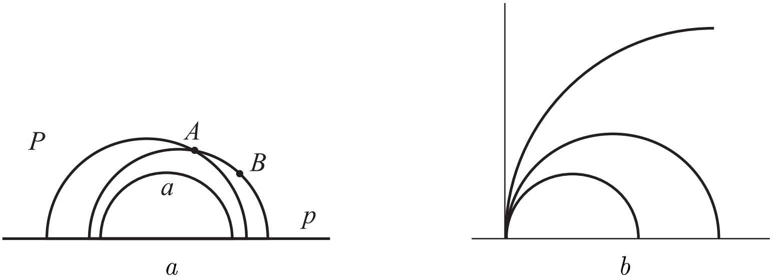


Рис. 32

двух «лучей» с общим началом, не содержащихся в одной «прямой» (рис. 32, *а*).

Помимо указанных «прямых» есть еще «прямые» — это полупрямые, перпендикулярные граничной прямой. Они являются пределами рассмотренных полуокружностей (рис. 32, *б*). Когда центр полуокружности удаляется по граничной прямой, а полуокружность проходит через данную точку, то она «распрямляется» и в пределе переходит в полупрямую. Поэтому мы дальше будем мыслить указанные полупрямые среди «прямых» модели в качестве полуокружностей, как «полуокружности бесконечного радиуса». Это позволит обойтись без скучных оговорок, касающихся этих полупрямых, причем, однако, следует помнить условность этого и быть готовым проверять утверждения для таких «полуокружностей». («Отрезок» на такой «прямой» — это обычный отрезок, а «лучи» — один обычный луч, другой — отрезок с исключенным концом на граничной прямой.)

Рассмотрим теперь в этой модели те аксиомы, в которые не входит понятие о равенстве отрезков и углов.

Аксиома параллельных для прямых относится к таким аксиомам. В данной модели она явно не выполняется: через точку A, не лежащую на «прямой» a, проходит бесконечно много «прямых», не имеющих с a общих точек (рис. 32, *а*).

Все прочие аксиомы, говорящие о связи точек и отрезков или точек и прямых, о взаимном расположении точек и прямых, здесь выполняются. Так, на рис. 33 указано построение отрезка с данными концами. Далее, возьмем полуокружность, представляющую «прямую» в модели. Проведем прямую l, касающуюся этой полуокружности и парал- лельную граничной прямой. Спроектируем полуокружность из ее центра на прямую l (рис. 34). Получим взаимно однозначное, сохра-

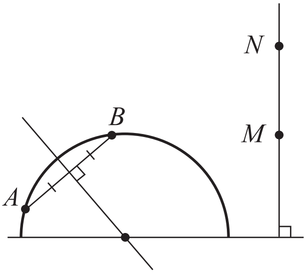


Рис. 33

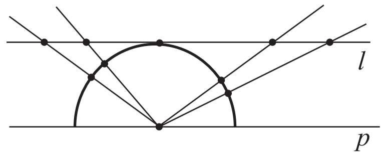


Рис. 34.

няющее порядок точек, соответствие между точками прямой и полуокружности, т. е. «прямой» модели. Все свойства, выраженные в аксиомах, будут одни и те же. Они также очевидно выполнены на полупрямых, представляющих «прямые» модели. Аксиома деления плоскости также выполняется. «Прямая» — полуокружность — делит плоскость на две области — внутреннюю и внешнюю. Это и будут «полуплоскости» в нашей модели. Из одной в другую нельзя перейти по какой-либо дуге, не пересекая разделяющую их «прямую» — полуокружность.

Остается определить равенство «отрезков» и «углов» так, чтобы выполнялись соответствующие аксиомы. Это мы сделаем, определив «наложение». Сначала определим «отражение в прямой». За «отражение в прямой» примем инверсию в той окружности, полуокружность которой представляет данная «прямая». Если же «прямая» — это полупрямая, перпендикулярная граничной прямой, то «отражением» в ней будет обычное отражение.

«Наложением» в модели называем любую композицию «отражений». «Равными» считаем фигуры, в частности, «отрезки» и «углы», совмещаемые «наложением».

Это определение сразу приводит к выводу: углы, «равные» в модели, равны без кавычек — в обычном смысле. В самом деле, углы при инверсиях сохраняются, т. е. преобразуются в равные, но они «равны» в модели по определению. Обратно: углы, «равные» в модели, — это те, которые преобразуются друг в друга «наложениями», т. е. инверсиями, и, стало быть, они равны в обычном смысле.

При инверсии в окружности с центром на граничной прямой эта прямая и полуплоскость P отображаются на себя. Поэтому содержа-

щаяся в P полуокружность с центром на граничной прямой отображается на такую же полуокружность. В модели это означает, что при «отражениях» «прямые» переходят в «прямые». Очевидно, что также «лучи» переходят в «лучи» и «отрезки» — в «отрезки».

Обратимся к откладыванию отрезков и углов в модели. Понятия, относящиеся к модели, будем предварять знаком *.

Пусть даны точка A, *луч a с началом A, *отрезок AB на этом *луче и *угол ab с вершиной A , образованный *лучом a вместе с *лучом b. Пусть даны также точка A', исходящий из нее *луч a', и отмечена *полуплоскость Q, ограниченная *прямой, содержащей *луч a' (рис. 35, a). Нам нужно произвести *наложение, переводящее точку A в A', *луч a — в a' и *луч b — в *луч, лежащий в *полуплоскости Q так, что *угол, *равный ab, отложится от a' в эту *полуплоскость.

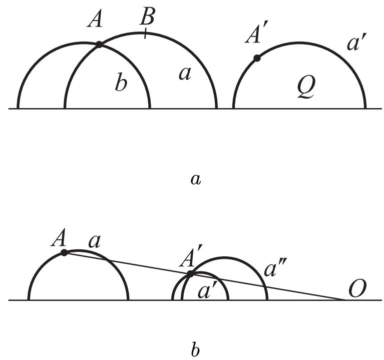


Рис. 35

Проведем прямую AA', и пусть она пересекает граничную прямую p в точке O (рис. 35, $б$). Произведем инверсию с центром O, которая переведет A в A'. *Луч a перейдет в *луч a'' с началом A', он образует с *лучом a' *угол $a'a''$.[46]

Проведем прямую q (без кавычек), делящую *угол $a'a''$ пополам, и построим окружность с центром на граничной прямой, касающуюся прямой q (кстати, укажите такое построение). Инверсия в этой окружности переведет *луч a'' в a' (почему?). В смысле модели это значит, что *отражение в соответствующей *прямой переводит *луч a'' в a'. Таким образом, два *отражения переводят точку A в A' и *луч a — в a'. Вместе с *лучом вся содержащая его *прямая $\overline{a}$ — полуокружность — переходит в *прямую $\overline{a}'$ — полуокружность, — содержащую *луч a. *Полуплоскости, ограниченные *прямой $\overline{a}$, отображаются на *полуплоскости, ограниченные *прямой $\overline{a}'$. *Луч b, служащий стороной данного *угла ab, переходит в *луч b'' с началом A'. Но он может оказаться не в той *полуплоскости, которая была заранее отмечена. Тогда нужно произвести еще *отражение в *прямой, содержащей *луч

[46] Если прямая AA' параллельна граничной прямой p, то точка A переводится в A' отражением в прямой, перпендикулярной p, т. е. *отражением в прямой, представляющей полупрямую, перпендикулярную p.

a', т. е. инверсию в окружности, содержащей эту *прямую. При этом на самой *прямой $\overline{a}'$ ничего не происходит: все ее точки остаются неподвижными. И только *луч b'' перейдет в *луч b, лежащий в указанной *полуплоскости.

Если на *луче a была отмечена какая-нибудь точка B, и тем самым отмечен *отрезок AB, то эта точка перейдет в определенную точку B' на *луче a' и *отрезок AB — в *отрезок $A'B'$ на этом *луче. Так мы получаем результат: на каждом *луче a' можно от его начала отложить *отрезок, *равный данному, т. е. для любого данного *отрезка AB на данном *луче с началом A' есть такая точка B', что *отрезок AB можно перевести в *отрезок $A'B'$ путем *наложения.

Совершенно так же то, что *луч b перейдет в *луч b', лежащий в нужной полуплоскости, что и *угол $a'b'$ равен данному ab, позволяет утверждать:

От каждого *луча от его начала по данную сторону от *прямой, его содержащей, можно отложить * угол, равный данному.

Остается доказать, что *угол откладывается единственным образом, так же, как и *отрезок (или, по нашей аксиоме меньшего отрезка, отрезок, содержащийся в данном и не совпадающий с ним, не может быть равен ему).

Утверждение о единственности откладывания угла сводится, очевидно, к следующему:

Если *лучи b, c, исходящие из начала *луча a, образуют с ним равные углы и лежат с одной стороны от него (в одной полуплоскости), то они совпадают.

Но *углы, равные в модели, равны в обычном «евклидовом» смысле, а для обычных углов сказанное, очевидно, верно. *Лучи b, c содержатся в окружностях с центрами на данной прямой p. Раз они образуют с *лучом a данный угол, то, значит, дана касательная к указанным окружностям в точке A. Но окружность с центром на данной прямой, касающаяся другой прямой в данной ее точке, только одна. Значит, *лучи b, c совпадают. Итак, *угол откладывается единственным образом.

*Отрезок, *равный данному, также откладывается на данном *луче единственным образом. Действительно, пусть *отрезок AB, *равный данному, отложен на данном *луче a с началом A. Если бы можно было отложить другой *отрезок, AC, равный тому же, то это значило бы, что есть *наложение (отличное от тождественного), отображающее *луч сам на себя. Оно отображает тогда на себя и всю содержащую его *прямую — полуокружность $\overline{a}$. Если же *наложение переставляет

*полуплоскости, ограниченные *прямой $\overline{a}$, то добавив отражение в ней, можно добиться того, что и полуплоскости эти будут отображаться каждая на себя.

В таком случае, ввиду сохранения углов, все *лучи, исходящие из точки A, будут отображаться на себя. Значит, при такой композиции инверсий (и отражений в вертикальных лучах) все концы лучей на граничной прямой остаются на месте. Вместе с ними отображаются на себя все полуокружности с концами на граничной прямой, т. е. *прямые модели. Но каждую точку можно получить в пересечении этих *прямых. Поэтому все точки отображаются на себя — «остаются на месте» — так что рассматриваемое *наложение оказывается тождественным вопреки предположению.

Этим единственность откладывания на данном луче отрезка, равного данному, доказана.

На этом доказательство того, что в рассмотренной модели выполняется геометрия Лобачевского, заканчивается. Требование аксиомы меньшего отрезка, что в отрезок нельзя уместить ему равный, заведомо выполняется при том, что уже доказано. Впрочем, доказательство того, что оно выполнено, читатель может провести сам.

Описанную модель плоскости Лобачевского можно еще назвать *конформной*, поскольку в ней наложения представляются инверсиями — преобразованиями, сохраняющими углы.

Модель геометрии Лобачевского в пространстве. Эта модель определяется аналогично модели на плоскости.

За пространство принимается открытое полупространство P. «Плоскостями» в нем служат содержащиеся в P полусферы с центрами на граничной плоскости, а также перпендикулярные ей открытые полуплоскости (рис. 36). За «прямые» принимаются полуокружности, перпендикулярные граничной плоскости (т. е. касательные к ним в концах перпендикуляры этой плоскости; центры их лежат на граничной плоскости), а также перпендикулярные ей лучи.

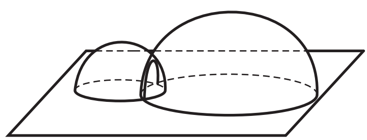


Рис. 36

Роль «наложений» играют композиции инверсий в сферах с центрами на граничной плоскости и отражений в перпендикулярных ей плоскостях.

Проективная модель геометрии Лобачевского. Геометрия Лобачевского на плоскости представляется еще следующей моделью на обычной евклидовой плоскости.

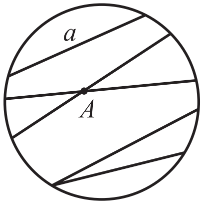


Рис. 37

Плоскость Лобачевского представляется внутренностью круга (рис. 37), прямые — хордами (с исключенными концами, поскольку рассматривается только внутренность круга).

Преобразования — отображения круга на себя, переводящие хорды в хорды, — принимаются за наложения; так что равными считаются фигуры внутри круга, которые отображаются одна на другую при таких преобразованиях круга.

Преимущество этой модели в том, что прямые изображаются очень наглядно: хордами — прямолинейными отрезками. Но зато в отличие от предыдущей модели углы в смысле геометрии Лобачевского связаны с обычными евклидовыми углами очень сложно.

Геометрия Лобачевского в пространстве представляется аналогичной моделью. Пространством служит внутренность шара, прямыми — хорды с исключенными концами, наложениями — отображения шара на себя, переводящие хорды в хорды. Плоскости представляются внутренностями кругов, являющихся плоскими сечениями шара (это уже очевидно следует из того, что прямые изображаются хордами).

Указанную модель для плоскости и для пространства мы назвали *проективной*, так как отображения, которые представляют здесь наложения, — это проективные преобразования соответственно круга и шара. Эта модель называется также *моделью Кэли — Клейна.* Ее построил английский математик Кэли, но он не понял, что введенная им геометрия в круге и есть геометрия Лобачевского; это сообразил позже, в 1870 г., немецкий математик Клейн.

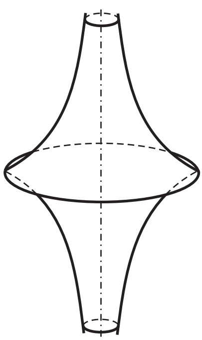


Рис. 38

Модель геометрии Лобачевского на поверхности. Оказывается, что геометрия Лобачевского реализуется на поверхностях постоянной отрицательной кривизной: внутренняя геометрия такой поверхности и есть геометрия Лобачевского. Только не на всей плоскости, а на той ее части, которая может быть представлена данной поверхностью. Вместе с тем доказано, что не существует (в трехмерном евклидовом пространстве) никакой поверхности, которая своей внутренней геометрией представляла бы всю плоскость Лобачевского.

Во внутренней геометрии поверхности роль прямолинейных отрезков играют кратчайшие линии (отрезки геодезических); роль наложений — такие отображения фигур, содержащихся в поверхности, которые сохраняют расстояния, измеряемые по этим кратчайшим линиям.

Самая известная из поверхностей постоянной отрицательной кривизны — псевдосфера — изображена на рис. 38.

Реализацию геометрии Лобачевского на поверхностях постоянной отрицательной кривизны установил итальянский математик Бельтрами (в 1861 г.). Впрочем, еще за 30 лет до него это установил, собственно, Миндинг — профессор университета в Дерпте (ныне Тарту), — но не понял этого.

§ 2. Факты геометрии Лобачевского

Аксиоматика геометрии Лобачевского полная — все ее модели изоморфны. Поэтому факты — теоремы этой геометрии можно наблюдать на любой модели.

Здесь мы укажем ряд фактов геометрии Лобачевского, которыми она отличается от геометрии Евклида, прежде всего на плоскости. Мы иллюстрируем их на конформной или проективной модели.

1. По аксиоме Лобачевского на плоскости через точку A, не лежащую на прямой a, проходит по крайней мере две прямые, не пересекающие a.

Если AB — перпендикуляр из A на прямую a, то прямая b, проходящая через A перпендикулярно AB (рис. 39, a), заведомо не пересекает a (что доказывается так же, как в евклидовой планиметрии). Поэтому в силу аксиомы Лобачевского есть прямая, проходящая через A и образующая с AB с одной стороны острый угол, но не пересекающая прямой a. Пусть α — точная нижняя граница таких углов. Прямая c, проходящая под таким углом к AB, все еще не пересекает a. (Так как иначе через A проходила бы прямая, пересекающая a в более далекой точке и, стало быть, образующая с AB больший угол.) В то же время любая прямая, проходящая через A и образующая с AB угол, меньший, чем α, уже пересекает a. Такая прямая c в геометрии Лобачевского называется *параллельной* прямой a. Острый угол α, который она образует с перпендикуляром AB, называется *углом параллельности*. По симметрии есть вторая прямая c', параллельная a, образующая с AB угол α с другой стороны. Эти прямые неограниченно, асимптотически приближаются к прямой a с одной и другой стороны, т. е. к одному ее лучу и к другому.

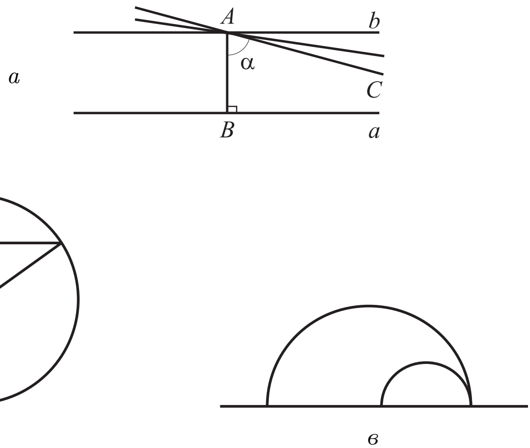


Рис. 39

В конформной модели эти прямые изображаются окружностями, касающимися окружности, изображающей прямую a, в точках на граничной прямой (рис. 39, *б*). В проективной модели параллельные прямые изображаются хордами с общим концом (рис. 39, *в*).

Угол параллельности α зависит от расстояния $h = |AB|$. С ростом h он убывает и стремится к нулю, как видно на конформной модели (рис. 40). Имеет место формула Лобачевского $\operatorname{tg} \alpha/2 = e^{-hk}$, где k зависит от единицы измерения длины.

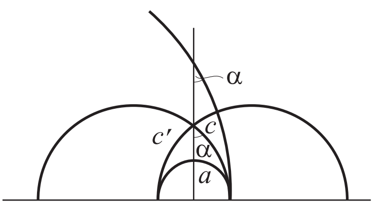


Рис. 40

2. В геометрии Лобачевского не выполняется наша аксиома параллельных отрезков: если равные отрезки AC и BD проведены перпендикулярно отрезку AB в одну сторону от него, то $CD \neq AB$. Более того, заведомо $CD > AB$, и с увеличением длин отрезков AC, BD расстояние CD неограниченно растет. Это значит, что прямые AC и BD с общим перпендикуляром AB неограниченно расходятся в обе стороны. Выполняется даже следующее.

Пусть прямая a перпендикулярна отрезку AB в его середине. Тогда прямые, ей перпендикулярные и пересекающие ее достаточно да-

леко от AB, не пересекают прямых AC, BD — лежат в полосе между ними.

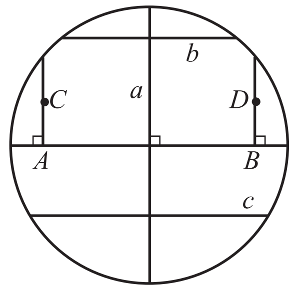
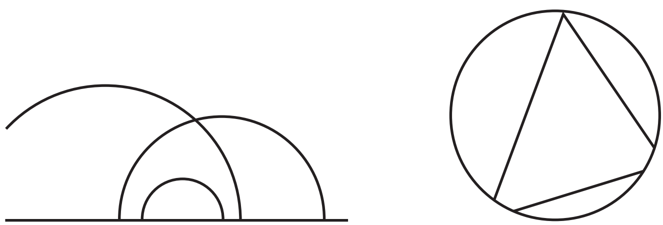


Рис. 41

Это видно на проективной модели (рис. 41). Из симметрии относительно диаметра круга ясно, что хорда, перпендикулярная диаметру, изображает прямую, перпендикулярную прямой, изображаемой диаметром. Поэтому на рисунке хорды b, c — это прямые, перпендикулярные a. Посмотрите, как это выражается в конформной модели.

3. Во всяком сколь угодно малом угле (как части плоскости) содержатся прямые, не пересекающие его сторон. В частности, среди таких прямых есть и прямые, перпендикулярные биссектрисе. Это так же видно на моделях (рис. 42).

4. Если две прямые не пересекаются, то они либо имеют, и притом единственный, общий перпендикуляр и бесконечно расходятся друг от друга в обе стороны, либо они параллельны и расходятся в одну сторону, а в другую — асимптотически сближаются.



а *б*

Рис.42

5. Из сказанного в 4 следует, что никакие две прямые не располагаются на постоянном расстоянии друг от друга. Линия, проходящая на постоянном расстоянии от прямой — выпуклая кривая; она называется *эквидистантой*. При сдвиге вдоль прямой точки плоскости, не лежащие на этой прямой, перемещаются по эквидистантам.

6. Сумма углов треугольника не равна π; она всегда меньше на величину, пропорциональную площади треугольника: $\alpha + \beta + \gamma = \pi - k^2 S$, где k зависит от единицы измерения длины.

7. Не существует треугольников сколь угодно большой площади. (Это уже следует из предыдущего: так как $\alpha+\beta+\gamma>0$, то $k^2s<\pi$, $s<\pi/k^2$.)

8. Нет подобных треугольников; если у двух треугольников углы равны, то треугольники равны. По этому существует «абсолютная» единица длины, определяемая самой геометрией, например, сторона равностороннего треугольника с суммой углов $\pi/2$.

9. Длина окружности l не пропорциональна радиусу r, а растет быстрее: $l=\pi/k(e^{kr}-e^{-kr})=2\pi\cdot 1/k\,\mathrm{sh}\,kr$, где k — постоянная, указанная выше (kr — безразмерная величина), и sh обозначает гиперболический синус.

10. Предел бесконечно растущих окружностей, касающихся данной прямой в данной точке, — не эта прямая, а особая кривая, называемая *орициклом* (*или предельной окружностью*).

11. Через три точки, не лежащие на одной прямой, проходит либо окружность, либо орицикл, либо зквидистанта.

Геометрия Лобачевского связана с евклидовой тем, что во всякой достаточно малой области на плоскости Лобачевского приближенно выполняется геометрия Евклида, тем точнее, чем меньше область. Если за такую область взять круг радиуса r, то относительное отклонение от линейных соотношений евклидовой геометрии будет порядка $(kr)^2$, как в формулах для суммы углов треугольника и длины окружности (убедитесь из разложения в ряд).

§ 3. Многомерное евклидово пространство

Многомерное евклидово пространство уже было рассмотрено в гл. V ч. 3. Оно определяется геометрическими аксиомами, аналогичными аксиомам стереометрии; эти аксиомы изложены также в § 6 гл. I. Но совершенно так же, как была сформулирована в § 2 гл. III «координатная аксиоматика» планиметрии и стереометрии, можно сформулировать координатную аксиоматику n-мерного евклидова пространства. Она состоит, можно сказать, в том, что сказанное в теоремах 1, 2 § 1 гл. V ч. 3 принимается в качестве аксиом. Итак, формулируем.

Координатная аксиоматика n-мерного евклидова пространства.

Основные ее объекты — *точки.*

Аксиомы.

1. *Каждой точке M соответствует определенная (единственная) последовательность n чисел $(x_1, x_2,\dots, x_n)$ — «координаты точ-*

ки M», так что каждая последовательность $(x_1, x_2, \ldots, x_n)$ соответствует какой-либо точке. Тем самым между точками и последовательностями (по n чисел) *есть взаимно однозначное соответствие.*

2. *Каждой паре точек $M(x_1, \ldots, x_n)$, $M'(x'_1, \ldots, x'_n)$ поставлено в соответствие число*

$$|MM'| = \sqrt{(x_1 - x'_1)^2 + \ldots + (x_n - x'_n)^2} \tag{1}$$

— численное расстояние между точками.

3. *Геометрическими считаются такие и только такие соотношения, которые определяются расстоянием между точками и сохраняются при умножении расстояний на один и тот же, но любой положительный множитель.*

С помощью расстояния определяются отрезки, так же как это сделано в случае планиметрии, т. е. при $n = 2$ (в § 2 гл. I). Равенство отрезков определяется равенством расстояний между концами. Вместе с отрезками определяются прямые (так же, как это сделано в § 7 гл. I о фигурах). Наконец, определяются плоскости, например так. Плоскость, проходящая через две пересекающиеся прямые a, b есть объединение всех прямых, пересекающих обе прямые a и b, но в разных точках, с добавлением точки пересечения.

Таким образом, все основные понятия ГА — геометрической аксиоматики, изложенной в § 1 гл. V ч. 3, — оказываются определенными с помощью расстояния. После этого доказывается, что при указанном их определении выполняется все сказанное в аксиомах ГА (доказательство мы не даем, оно будет намечено в конце этого параграфа). С другой стороны, аксиомы КА — координатной аксиоматики — следуют из ГА, как теоремы 1, 2 § 1 гл. V ч. 3. Таким образом, обе системы аксиом — геометрические и координатные — равносильны.

В КА основными — заранее не определимыми — объектами являются только точки. Если же точку отождествить с набором ее координат, то ничего не определяемого не останется и мы получим числовую модель аксиоматики n-мерного евклидова пространства. Наличие такой модели доказывает непротиворечивость этой аксиоматики и в этом смысле — существование n-мерного евклидова пространства в качестве непротиворечиво мыслимого объекта (в меру того, насколько непротиворечива теория вещественных чисел, из которых модель строится).

Введение координат в n-мерном евклидовом пространстве доказывает, так же как в случае планиметрии, что все возможные модели

этого пространства (с данным n) изоморфны координатной модели. Тем самым аксиоматика n-мерного евклидова пространства полная.

Замечание. Исторически понятие о многомерном пространстве возникло из той идеи, что можно мыслить пространство, в котором точка определяется не тремя, а бóльшим числом координат. Соответственно, координатное определение n-мерного пространства было первым, и вопрос о его непротиворечивости не стоял.

Групповое определение евклидовой геометрии. Из того, что соотношения евклидовой геометрии выражаются через расстояния, взятые с точностью до общего множителя, следует, что все эти соотношения сохраняются при любых отображениях, при которых расстояния только умножаются на одно и то же число. С другой стороны, если некоторое соотношение сохраняется при любых таких отображениях, сохраняющих расстояния с точностью до множителя, то оно от них только и зависит.

Таким образом, можно сказать, что евклидовой геометрии принадлежат все те и только те соотношения, какие сохраняются при указанных преобразованиях, — при которых все расстояния только умножаются на общий множитель, или, что равносильно, сохраняются отношения расстояний. Такие преобразования пространства образуют группу (уже в связи с наложениями было отмечено, что обратимые отображения любого множества на себя, что-либо в нем сохраняющие, образуют группу). В планиметрии и стереометрии это преобразования подобия. Такое название переносится в n-мерную геометрию. Поэтому можно сказать: *евклидова геометрия любого числа измерений имеет своим предметом свойства фигур, сохраняющиеся при подобных преобразованиях или, как говорят, инвариантные относительно группы подобных преобразований.*

По первой координатной аксиоме каждой точке отнесен набор чисел — ее координат $x_1, x_2, \ldots, x_n$, и точки различаются координатами. Преобразование сопоставляет точкам $(x_1, \ldots, x_n)$ точки $(y_1, \ldots, y_n)$. Оно, стало быть, изображается преобразованием координат:

$$y_1 = f_1(x_1, \ldots, x_n), \ \ldots, \ y_n = f_n(x_1, \ldots, x_n). \qquad (2)$$

Условие состоит в том, что это преобразование может придавать расстоянию только множитель. Вместо расстояния можно, конечно, взять подкоренное выражение

$$(x_1 - x_1')^2 + \ldots + (x_n - x_n')^2, \qquad (3)$$

и тогда условие приводится к тому, что должно быть

$$(y_1 - y_1')^2 + \ldots + (y_n - y_n')^2 = = \lambda[(x_1 - x_1')^2 + \ldots + (x_n - x_n')^2], \quad (4)$$

где λ зависит не от координат, а только от самого преобразования.

Теперь мы можем сформулировать:

Групповые аксиомы евклидова пространства.

1. *Точки находятся во взаимно однозначном соответствии с наборами чисел* $x_1, \ldots, x_n$ *— их координат.*

2. *Геометрическим считается всякое выраженное в координатах соотношение, сохраняющееся при всех тех преобразованиях* (2), *при которых выполняется равенство* (4).

Другими словами, геометрические — это те соотношения, которые инвариантны относительно группы подобий — преобразований (2) с условием (4).

Какие это преобразования — рассказывается в курсе алгебры. Так же как на плоскости и в трехмерном пространстве, это — ортогональные преобразования с добавлением умножения на любое число, отличное от нуля. То есть преобразования вида

$$\begin{aligned} &y_1 = \lambda(a_{11}x_1 + \ldots + a_{1n}x_n) + b_1, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &y_n = \lambda(a_{n1}x_1 + \ldots + a_{nn}x_n) + b_n \end{aligned} \qquad (5)$$

с условиями: сумма квадратов коэффициентов a_{ik} в каждой строке равна единице, а суммы произведений одноименных коэффициентов из разных строк равны нулю:

$$\begin{aligned} &a_{k1}^2 + \ldots + a_{kn}^2 = 1 \quad (k = 1,\ \ldots,\ n), \\ &a_{i1}a_{k1} + \ldots + a_{in}a_{kn} = 0 \quad (i \neq k;\ i,\ k = 1,\ \ldots,\ n). \end{aligned} \qquad (6)$$

Аксиому 2 можно формулировать так:

2. Евклидовой геометрии принадлежат те и только те соотношения, которые инвариантны относительно группы подобий, т. е. преобразований вида (5) с условиями (6).

План построения геометрии на координатных аксиомах. Рассмотрим n-мерное евклидово пространство с координатными аксиомами. Каждая точка определяется своими координатами $(x_1, \ldots, x_n)$.

Подобно тому, как это можно сделать в стереометрии, определим конкретные векторы как упорядоченные пары точек: A — начало, B —

конец вектора $\overrightarrow{AB}$. Координаты вектора определяются как разности координат конца и начала: $x_1^B - x_1^A$ и т. д. Векторы равны, если равны их координаты: равные векторы — пары точек — являются представителями одного «абстрактного» вектора, который полностью характеризуется координатами: $\boldsymbol{a} = \langle a_1, \ldots, a_n \rangle$. (Не путать векторы с точками!)

Сложение векторов, умножение на число и скалярное произведение определяем формулами

$$\boldsymbol{a} + \boldsymbol{b} = \langle a_1 + b_1, \ \ldots, \ a_n + b_n \rangle,$$
$$\lambda \boldsymbol{a} = \langle \lambda a_1, \ \ldots, \ \lambda a_n \rangle, \ \boldsymbol{ab} = a_1 b_1 + \ \ldots + a_n b_n.$$

Из последнего равенства следует, если положить $\boldsymbol{aa} = \boldsymbol{a}^2$ и т. п.,

$$\boldsymbol{ab} = \frac{1}{2}((\boldsymbol{a} + \boldsymbol{b})^2 - \boldsymbol{a}^2 - \boldsymbol{b}^2).$$

Тем самым скалярное произведение выражено через расстояние.

Если $x_1, \ \ldots, \ x_n$ — координаты точки M, то с ней связываем ее «радиус-вектор» с ее координатами $\boldsymbol{r} = \langle x_1, \ \ldots, \ x_n \rangle$.

Прямая, проходящая через точку с радиус-вектором $\boldsymbol{x}_0$, есть множество точек с радиус-векторами

$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{a}t, \quad -\infty < t < +\infty,$$

$\boldsymbol{a}$ — направляющий вектор.

Плоскость можно задать ее точкой x_0 и двумя неколлинеарными друг другу векторами $\boldsymbol{a}$, $\boldsymbol{b}$; точки плоскости имеют радиус-векторы

$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{a}t + \boldsymbol{b}s, \quad -\infty < t, \ s < \infty. \tag{7}$$

По векторам $\boldsymbol{a}$, $\boldsymbol{b}$ можно определить взаимно перпендикулярные единичные векторы

$$\boldsymbol{e}_1 = \frac{\boldsymbol{a}}{|\boldsymbol{a}|}, \quad \boldsymbol{e}_2 = \frac{\boldsymbol{b} - \boldsymbol{e}_1(\boldsymbol{e}_1\boldsymbol{b})}{|\boldsymbol{b} - \boldsymbol{e}_1(\boldsymbol{e}_1\boldsymbol{b})|}.$$

Выражая в (7) $\boldsymbol{a}$ и $\boldsymbol{b}$ через $\boldsymbol{e}_1$ и $\boldsymbol{e}_2$, получим

$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{e}_1 t_1 + \boldsymbol{e}_2 t_2.$$

Этим на плоскости введены прямоугольные координаты t_1, t_2 (с основными векторами $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ и с началом в точке x_0). С помощью этих координат убеждаемся, что на плоскости выполняется планиметрия.

Теперь уже нетрудно проверить, что в пространстве, определенном координатными аксиомами, выполняются геометрические аксиомы.

§ 4. Групповой принцип оснований геометрии

Наряду с евклидовой геометрией, которая составляет наш главный предмет, во II и III частях и настоящей главе мы встретились, хотя и кратко, со сферической геометрией, с аффинной геометрией, с проективной и круговой геометриями и с геометрией Лобачевского. В общем — шесть разных «геометрий», причем все они обобщаются на произвольное число измерений. Однако при всем разнообразии их объединяет общий принцип, который может быть положен в их основание.

Каждая из этих геометрий имеет своим предметом свойства фигур, сохраняющиеся при тех или иных преобразованиях. Эти преобразования пространства образуют соответствующую группу, так что общее основание указанных систем геометрии можно выразить так: *в каждой из них изучаются свойства фигур, инвариантные относительно преобразований соответствующей группы.*

В евклидовой геометрии изучаются свойства фигур, сохраняющиеся при подобных преобразованиях, в сферической — при поворотах сферы. Предмет аффинной геометрии составляют свойства, сохраняющиеся при аффинных преобразованиях, так же как в проективной геометрии изучаются свойства, сохраняющиеся при проективных преобразованиях; в круговой геометрии — свойства, сохраняющиеся при преобразованиях, порождаемых инверсиями, включая отражения в плоскостях (как инверсии в «сферах бесконечного радиуса»). Наконец, к геометрии Лобачевского относятся те свойства фигур, которые сохраняются при наложениях в смысле этой геометрии; они были представлены в модели с помощью инверсий (§ § 1, 2).

Во второй части предыдущего параграфа было дано формальное определение евклидовой геометрии произвольного числа измерений с точки зрения указанного группового принципа. Аналогично можно определить аффинную геометрию n измерений двумя условиями — аксиомами.

Основные ее объекты — точки. Аксиомы следующие.

1. *Точки находятся во взаимно однозначном соответствии с наборами чисел* $x_1, \ldots, x_n$ *— их координат.*

2. *Аффинной геометрии принадлежат те и только те соотношения, которые сохраняются при любых преобразованиях, представляющихся в координатах как линейные, т. е. если точка* $(x_1, \ldots, x_n)$ *переходит в точку* $(y_1, \ldots, y_n)$, *то*

$$y_1 = a_{11}x_1 + \ldots + a_{1n}x_n + b_1,$$

$$y_n = a_{n1}x_1 + \ldots + a_{nn}x_n + b_n.$$

Поскольку преобразование по понятию обратимо, то эти формулы должны быть обратимы — так что определитель из коэффициентов не равен нулю. Эти преобразования образуют «аффинную группу», и предмет аффинной геометрии составляют ее инварианты.

Аналогично можно определить проективную геометрию n измерений, пользуясь однородными координатами. Точка определяется $n+1$ однородными координатами $x_1, \ldots, x_{n+1}$. Преобразование изображается однородными линейными формулами с определителем, отличным от нуля, как записано для $n = 2$ в § 4 гл. III.

Эти преобразования образуют «проективную группу» n измерений, и предмет проективной геометрии составляют ее инварианты.

Представлять подобным образом в координатах круговую геометрию и геометрию Лобачевского мы здесь не будем.

Отметим, что для того, чтобы инверсии «действовали» в пространстве (как и в плоскости) взаимно однозначно, его нужно дополнить бесконечно удаленной точкой. Поэтому преобразования, порождаемые инверсиями, заведомо нельзя представить в обычных координатах (как и проективные преобразования изображаются не в обычных, а в однородных координатах).

§ 5. Геометрия теории относительности

Наряду с евклидовой геометрией рассматривают «псевдоевклидову геометрию». В координатах, когда точка определяется набором координат, с точки зрения группового принципа, изложенного в предыдущем параграфе, это выглядит так.

Предмет евклидовой геометрии составляют инварианты преобразований, сохраняющих, с точностью до множителя, сумму квадратов разностей координат двух точек, т. е. обозначая разность буквой Δ — дельта:

$$(\Delta x_1)^2 + (\Delta x_2)^2 + \ldots + (\Delta x_n)^2. \tag{1}$$

Предмет же псевдоевклидовой геометрии n изменений составляют инварианты преобразований, которые сохраняют, с точностью до множителя, величину

$$V = (\Delta x_1)^2 + \ldots + (\Delta x_{n-1})^2 - (\Delta x_n)^2, \tag{2}$$

или вообще такую величину V, в выражение которой входит некоторое число m квадратов со знаком минус. Соответственно, псевдоевклидова геометрия может быть в разных вариантах по числу отрицательных квадратов в выражении, которое сохраняется с точностью до множителя[47].

При числе изменений $n > 2$, т. е. исключая случай плоскости, это оказывается равносильно сохранению равенства $V = 0$. В частности, при одном отрицательном квадрате, как в формуле (2), это означает сохранение равенства

$$(\Delta x_1)^2 + \ldots + (\Delta x_{n-1})^2 - (\Delta x_n)^2 = 0. \tag{3}$$

Ограничимся этим случаем с одним отрицательным квадратом при числе измерений $n = 4$ (почему так — будет ясно дальше). Соответствующую псевдоевклидову геометрию можно определить так.

Основные ее объекты — *точки.*

Аксиомы или определяющие условия:

1. *Каждая точка задается упорядоченным набором четырех чисел — координат x, y, z, t так, что между точками и всевозможными такими наборами есть взаимно однозначное соответствие.*

2. *Геометрический смысл имеют те и только те соотношения, которые сохраняются при любых преобразованиях, не нарушающих равенства*

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (\Delta t)^2 = 0, \tag{4}$$

для любых двух точек, для которых оно выполняется.

Геометрия теории относительности. Обратимся от отвлеченно мысленных отношений геометрии к реальной действительности, как мы можем ее себе представить в согласии с физикой.

Мир представляет собой множество событий, а событие характеризуется местом и временем, где и когда оно происходит. Место определяется тремя координатами x, y, z, время события определяется по промежутку времени t, прошедшему от некоторого начального момента. Таким образом, каждому событию соответствует свой набор четы-

[47]Преобразование, сохраняющее какую-либо величину P, сохраняет, очевидно, и P. Поэтому преобразование, сохраняющее с точностью до множителя величину V с m отрицательными квадратами, сохраняет и величину $-V$ с $n - m$ отрицательными квадратами. Поэтому псевдоевклидова геометрия, соответствующая обоим случаям, — одна и та же. Соответственно, число возможных псевдоевклидовых геометрий n измерений равно $n/2$ при четном n и $(n-1)/2$ — при нечетном n.

рех чисел x, y, z, t — четыре координаты, три пространственные x, y, z и одна временная t. Мир имеет четыре измерения.

Координаты x, y, z мы представляем как прямоугольные, так что расстояние r между двумя местами — точками (x_1, y_1, z_1), (x_2, y_2, z_2) — выражается формулой

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}. \tag{5}$$

От одних событий к другим, от одного места к другому распространяется свет, и вообще какое бы то ни было электромагнитное излучение — оно же поток фотонов. Для краткости будем говорить, как это принято, о свете.

Если скорость света обозначить c, то расстояние r, пройденное светом от момента t_0 до момента t, будет

$$r = c(t - t_0). \tag{6}$$

Поэтому закон распространения света от события (x_0, y_0, z_0, t_0) — от вспышки в месте (x_0, y_0, z_0) в момент времени t_0, представляется в силу формулы (5) так:

$$\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} = c(t - t_0). \tag{7}$$

Возводя в квадрат и обозначая разности буквой Δ, получим закон распространения света в виде

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2 = 0. \tag{8}$$

Во всех этих выводах имеются в виду такие пространственно-временные координаты x, y, z, t, в которых распространение света представляется как происходящее с постоянной скоростью во всех направлениях. Разные системы координат могут быть связаны с разными телами, движущимися различным образом, и распространение света относительно них может выглядеть по-разному, подобно тому, как по-разному выглядит движение одних тел относительно других, как мимо пассажиров «плывут» дома.

Но можно считать твердо установленным, что системы координат, в которых закон распространения света представляется формулой (7) или (8), существуют[48]. Назовем их *лоренцевыми* по имени физика Г. Лоренца, нашедшего формулы преобразования таких координат.

---

[48]Понятно, как для каждого закона физики, сказанное верно лишь приближенно; однако с большой степенью точности.

Эти координаты связаны с так называемыми *инерциальными системами*, по отношению к которым выполняется «закон инерции»: тело, не испытывающее воздействий, движется относительно этой системы прямолинейно и равномерно.

Для инерциальных систем выполняется фундаментальный закон природы — принцип относительности: *по отношению к таким системам явления одной и той же природы протекают одинаково.* (Как для пассажиров в равномерно летящем самолете все происходит в самолете так, как если бы он стоял на земле.)

Но не вдаваясь в эти выводы физики, можно высказать принцип относительности просто в его математическом выражении:

Законы физики выражаются одинаково во всех лоренцевых системах координат. Эти координаты характеризуются тем, что в них распространение света представляется формулой (8). Поэтому принцип относительности можно выразить так:

Выражения законов природы инвариантны при преобразованиях одной системы лоренцевых координат в другую, т. е. по отношению к таким преобразованиям, которые оставляют неизменной формулу закона распространения света (7).

Этот же закон выражается формулой (8), а если выбрать единицы измерения так, чтобы скорость света была $c = 1$, то формула (8) приобретает вид:

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (\Delta t)^2 = 0. \tag{9}$$

Теперь подведем итог.

1. *Мир есть множество событий, и каждое событие задается четырьмя координатами* x, y, z, t, при чем существуют такие координаты — «лоренцевы», в которых закон распространения света выражается формулой (9).

2. *Выполняется принцип относительности, как он только что был сформулирован.*

Сравнив это с определением псевдоевклидовой геометрии, связанной с формулой (4), можно видеть полное совпадение.

В геометрии — в структуре мира основные объекты — это события, подобно точкам.

Законы этой структуры, подобно аксиомам геометрии, следующие.

1. *Каждому событию можно сопоставить четыре координаты* x, y, z, t, *так что между событиями и наборами координат устанавливается взаимно однозначное соответствие.*

2. *Объективный физический смысл имеют только те соотношения, которые сохраняются при любых преобразованиях, не нарушающих равенства* (9).[49]

Понятно, в законы физики входят разные физические величины помимо координат x, y, z, t, и они должны тоже соответствующим образом преобразовываться. Но ограничимся тем, что может быть выражено в одних координатах (как например, какой-нибудь закон движения материальной точки, когда ее координаты x, y, z оказываются функциями t). В координатах выражаются пространственно-временные соотношения. И вывод, который мы получили, состоит в том, что общая структура таких соотношений — пространственно-временная структура мира — есть структура, или, другими словами, геометрия четырехмерного псевдоевклидова пространства.

§ 6. Риманова геометрия и другие

Риманова геометрия, или, что то же, геометрия римановых пространств» представляет собой обобщение внутренней геометрии поверхностей на произвольное число измерений и так же относится к n-мерной евклидовой геометрии, как внутренняя геометрия поверхностей — к геометрии на плоскости. Она названа по имени ее создателя Римана[50].

Внутренняя геометрия поверхности определяется длинами кривых на ней, поэтому с точки зрения своей внутренней геометрии поверхность — это двумерное многообразие, в котором определены длины кривых, причем длина гладкой кривой выражается интегралом $\int ds$ от «линейного элемента» ds. Это соответствует тому, что длины измеряются как бы «бесконечно малыми шагами» ds — сумма их вдоль кривой и дает ее длину (так пояснял сам Риман).

На гладкой поверхности во введенных на ней координатах квадрат

---

[49] Эти преобразования можно назвать «общими преобразованиями Лоренца», в отличие от специального их вида, к которому обычно применяют название «преобразования Лоренца».

[50] Бернхарт Риман (1826–1866) — немецкий математик, один из величайших математиков XIX в.; ему принадлежит ряд фундаментальных идей и результатов в разных областях математики. В 1854 г. он прочел в Гёттингенском университете лекцию «О гипотезах, лежащих в основании геометрии». В ней он ввел общее понятие о пространстве в математике, включая функциональные пространства, указал основы общей теории пространств с измерением длин, включающей риманову геометрию, и указал на ее возможную связь с физикой. Лекция эта была опубликована только после его смерти, в 1868 г.

линейного элемента ds^2 выражается как квадратичная форма от дифференциалов координат — первая квадратичная форма поверхности.

Чтобы определить, что такое риманово пространство, достаточно перенести сказанное на n измерений.

Риманово пространство — это многообразие, в котором определены длины кривых, причем длина кривой выражается интегралом $\int ds$ от «линейного элемента» ds. Это соответствует тому, что длины измеряются как бы «бесконечно малыми шагами» ds; сумма их вдоль кривой и дает ее длину $s = \int ds$.

В координатах $x_1, \ldots, x_n$, введенных в многообразии или в его области, квадрат линейного элемента ds^2 выражается как квадратичная форма от дифференциалов координат:

$$ds^2 = \sum_{i,k=1}^{n} g_{ik} dx_i dx_k, \tag{1}$$

где коэффициенты g_{ik} являются непрерывными функциями точки — ее координат $x_1, \ldots, x_n$; при этом $g_{ik} = g_{ki}$. Форма (1) положительная, т. е. принимает только положительные значения, кроме того случая, когда $dx_1 = \ldots = dx_n = 0$.

Задание ds^2 формулой (1) значит следующее. Если в многообразии или в некоторой его области, где расположена кривая, введены координаты $x_1, \ldots, x_n$, то кривая может быть задана параметрическими уравнениями

$$x_1 = f_1(t), \ldots, x_n = f_n(t); \quad t \in [t_0,\ t_1].$$

Предполагается, что эти функции имеют непрерывные производные. Тогда, приняв для этих производных обозначение $d\boldsymbol{f}_i/dt = \dot{x}_i$, можно из (1) получить выражение для элемента длины ds данной кривой:

$$ds = \sqrt{\sum_{i,k=1}^{n} g_{ik} \dot{x}_i \dot{x}_k}\, dt,$$

и длина кривой представится интегралом:

$$s = \int ds = \int_{t_0}^{t_1} \sqrt{\sum_{i,k=1}^{n} g_{ik} \dot{x}_i \dot{x}_k}\, dt.$$

Так как коэффициенты g_{ik} суть непрерывные функции координат x_1, …, x_n, а на кривой $x_1 = f_1(t)$ и т. д., то g_{ik} оказываются непрерывными функциями t, так же как $\dot{x}_1$, …, $\dot{x}_n$. Поэтому интеграл имеет обычный смысл и существует[51].

Так же как на поверхности, расстояния в римановом пространстве между двумя точками M, N можно определить как точную нижнюю границу длин кривых $\widehat{MN}$, соединяющих эти точки. Это будет длина кратчайшей из кривых $\widehat{MN}$, если кратчайшая существует. Во всяком случае у каждой точки есть такая окрестность, что любые две ее точки можно соединить кратчайшей. Кратчайшие играют в римановой геометрии роль отрезков. Кривая, которая оказывается кратчайшей на каждом достаточно малом участке, называется *геодезической*, так же как на поверхностях. Но большие дуги геодезических могут и не быть кратчайшими, как, например, дуги больших кругов на сфере, бóльшие полуокружности.

В малой окрестности любой точки геометрия риманова пространства мало отличается от евклидовой: тем меньше, чем меньше окрестность. Это основано на том, что преобразованием координат можно привести квадратичную форму $\sum g_{ik} dx_i dx_k$ в данной точке к виду $\sum dx_i^2$, т. е. к квадрату линейного элемента евклидова пространства в прямоугольных координатах. Вблизи данной точки линейный элемент ds будет поэтому мало отличаться от евклидова $\sqrt{\sum dx_i^2}$. Поэтому и геометрия в окрестности будет близка к евклидовой.

Точнее это можно выразить следующим образом.

В n-мерном римановом пространстве в окрестности любой данной точки O можно ввести такие координаты x_1, …, x_n, что расстояние $|MN|$ между любыми двумя точками M, N этой окрестности можно представить формулой

$$|MN| = \sqrt{(\Delta x_1)^2 + \ldots + (\Delta x_n)^2} + \varepsilon, \tag{2}$$

где Δx_1, …, Δx_n — разности координат точек M, N, а ε таково, что $\varepsilon/|MN| \to 0$, когда точки M, N приближаются к O.

---

[51] Риман ввел понятие о более общих пространствах, в которых также определены длины кривых, выражающиеся интегралом $\int ds$ от линейного элемента ds. Но ds представляется не обязательно корнем из положительной квадратичной формы от дифференциалов координат dx_1, …, dx_n, а любой их положительной функцией, зависящей также от точки — от самих координат, $ds = f(x_i; dx_i) = f(x_1, \ldots, x_n; dx_1, dx_n)$, причем эта функция однородная относительно dx_1, …, dx_n, т. е. $f(x_i; \lambda dx_i) = |\lambda| f(x_i; dx_i)$, что и позволяет представить длину интегралом.

Начало координат можно взять в точке O и обеспечить, что кратчайшие — геодезические линии, проходящие через нее, — будут представляться как прямые уравнениями вида

$$x_1 = a_1 s, \ \ldots, \ x_n = a_n s. \tag{3}$$

Длина отрезка такой линии будет точно представляться евклидовой или пифагоровой формулой:

$$|MN| = \sqrt{\Delta x_1^2 + \ldots + \Delta x_n^2}. \tag{4}$$

В частности, $|OM|^2 = x_1^2 + \ldots + x_n^2$.

Это позволяет определить угол — величину угла между двумя линиями, исходящими из точки O, так же как в евклидовом пространстве. А так как за O можно принять любую точку, то угол определяется между любыми двумя линиями, исходящими из одной точки. Точно так же можно ввести другие понятия дифференциальной геометрии, поскольку они относятся к свойствам кривых и поверхностей «в точке», т. е. определяются их свойствами в сколь угодно малой ее окрестности (и поскольку отклонения от евклидовой геометрий не играют роли).

Но риманова геометрия вообще отличается от евклидовой и в малых областях, хотя и на величины более высокого порядка малости. По аналогии с поверхностями, у которых гауссова кривизна относится к их внутренней геометрии, говорят о «римановой кривизне», характеризующей отклонение геометрии данного риманова пространства вблизи данной его точки от евклидовой геометрии. Взятое в точном смысле, понятие кривизны риманова пространства сложно. Оно, так же как на поверхностях, связано с отличием от π суммы углов треугольника, сторонами которого служат кратчайшие линии.

О других «геометриях». При групповом принципе определения той или иной геометрии она вводится на пространстве в целом, и так определяется, при данном числе измерений, ряд пространств: евклидовы, аффинные и др. Но в римановых пространствах геометрия определяется в бесконечно малых областях, можно сказать, дифференциально, и распространяется интегрированием. Поэтому разных римановых пространств любого данного числа измерений — целый континуум.

Этот «дифференциальный» принцип применяется также и в других случаях, порождая новые континуумы разных пространств.

Как геометрия риманова пространства является евклидовой в малых областях с точностью до величин высшего порядка малости, так

можно рассматривать пространства, в которых в малых областях, тоже с точностью до величин высшего порядка, выполняется какая-либо другая геометрия, например аффинная, — это в пространствах так называемой аффинной связности.

Особое значение имеют псевдоримановы пространства — такие, в малых областях которых с точностью до малых величин высшего порядка выполняется псевдоевклидова геометрия.

В применении к физике эта геометрия четырех измерений служит математическим аппаратом теории всемирного тяготения, носящей название общей теории относительности. Эта теория так же обобщает теорию относительности, как псевдориманова геометрия обобщает псевдоевклидову геометрию. Пространство и время можно разделять только в ограниченных пределах, вместе же они образуют единую форму существования материи — пространство-время, и геометрия его псевдориманова. «Кривизна» его проявляется во всемирном тяготении — в этом, можно сказать, и состоит сущность объяснения всемирного тяготения в общей теории относительности.

Здесь геометрия демонстрирует свое чрезвычайное значение в понимании самых фундаментальных законов мироздания, так же как новейшие геометрические теории находят применение в физике элементарных частиц.

В общем, современная геометрия включает многочисленные теории разного рода пространств и фигур в них, и далеко не все были здесь упомянуты. Но их объединяет общая основа, если не считать некоторых не очень значительных теорий.

Всякое пространство, которое определяют и исследуют в геометрии, представляет собой, за редким исключением, топологическое пространство, наделенное той или иной дополнительной структурой — метрикой, группой преобразований и т. д. и т. д. Но при всем их разнообразии все они выросли из зерна обычного трехмерного евклидова пространства, из его углубленного изучения и обобщения, из связанных с ним наглядных представлений как, скажем, преобразование фигур перемещением или проектированием с плоскости на плоскость, обобщаемое в общем групповом принципе определения разных «геометрий», или измерение длин малым масштабом, которое легло в основу римановой геометрии и ее обобщений.